version: Monday 2$^{\text{nd}}$ April, 2018 @ 10:22

PHY2404S (2018)
Quantum Field Theory II

Prof. A.W. Peet

Department of Physics,
University of Toronto,
Toronto, Ontario,
Canada M5S 1A7.

These are the lecture notes for PHY2404S (2018) "Quantum Field Theory II".

The content in this set of notes is partly original and partly follows discussions in QFT textbooks by: Matthew Schwartz, Michael Peskin and Daniel Schroeder, Thomas Banks, Anthony Zee, Lewis Ryder, Pierre Ramond, and the encyclopaedic Steven Weinberg (3 volumes).

Contents

1 The Power of Symmetry

1.1 Noether's Theorem

Symmetry has proven to be an extremely powerful way of organizing our physics thoughts about Nature. You are probably already familiar with 10 spacetime symmetries from study of relativity: 4 spacetime translations, 3 spatial rotations, and 3 Lorentz boosts. Conservation laws associated to them ensure both linear and angular momentum conservation as well as sensible centre of mass motion. The counting goes similarly in other spacetime dimensions, except that in D dimensions there are D translation parameters, $d = D-1$ boost parameters, and $d(d-1)/2$ rotation parameters.

Other symmetries, such as the $U(1)$ gauge symmetry of electromagnetism, act on the *fields* rather than on the coordinates. Field space, as distinct from spacetime, is usually referred to as the **internal space** for the field, rather than the **external space** of the coordinates. The charge carried by a field can be thought of as like a handle pointing in field space, onto which a gauge boson can grab.

Gauge fields can be in three distinct phases of physical behaviour. The most familiar from our undergraduate work is the **Coulomb phase**, resulting in an inverse-square law in four spacetime dimensions as per intuition. Alternatively, like for QCD at low energy, the gauge field can be in a **confined phase**. The third possibility is the one we will explore in the next chapter, and is known as the **Higgs phase** with spontaneous symmetry breaking.

A very important theorem proved by Emmy Noether[1] says that every continuous symmetry gives rise to a conservation law. To see how to prove this, let us consider a general symmetry transformation which may act on both fields and coordinates. For the coordinates, we write

$$x^\mu \to x^{\mu\prime} = x^\mu + \delta x^\mu \,. \tag{1}$$

Internal symmetries act directly on the fields, which are said to carry a **representation** of the symmetry group. So let us consider a general collection of fields $\{\phi^a\}$, where a represents a general index which might or might not have anything to do with spacetime indices. Under a symmetry transformation, the total variation of the field is

$$\begin{aligned}
\Delta\phi^a &= \phi^{a\prime}(x') - \phi^a(x) \\
&= \phi^{a\prime}(x') + [-\phi^{a\prime}(x) + \phi^{a\prime}(x)] - \phi^a(x) \\
&= [\phi^{a\prime}(x') - \phi^{a\prime}(x)] + [\phi^{a\prime}(x) - \phi^a(x)] \\
&= (\partial_\mu\phi^a)\delta x^\mu + \delta\phi^a(x) \,,
\end{aligned} \tag{2}$$

where in the last step we used the chain rule for differentiation. Notice that we have obtained the second term from the direct functional variation of the field while the first, known as the transport term, arose more indirectly through the dependence of the field on coordinates. This is why we distinguish notationally between field variations $\delta\phi$ and total variations $\Delta\phi$.

To prove Noether's theorem, consider the action for the fields $\{\phi^a\}$:

$$S = \int d^D x \, \mathscr{L}[\phi^a] \,. \tag{3}$$

[1] Woohoo! a woman!

Then

$$\delta S = \int \left[\delta(d^D x)\mathscr{L}[\phi^a] + d^D x(\delta\mathscr{L})\right] . \tag{4}$$

How does the measure of integration vary under such a symmetry transformation? Let us define the Jacobian $J(x'|x)$: $d^D x' = d^D x J(x'|x)$. We have

$$\frac{\partial x^{\mu\prime}}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}\left\{x^\mu + \delta x^\mu\right\} = \delta^\mu_\nu + \partial_\nu(\delta x^\mu) . \tag{5}$$

Therefore, to first order in small quantities,

$$J(x'|x) = 1 + \partial_\mu(\delta x^\mu) + \dots . \tag{6}$$

How about variations in the Lagrangian? Assuming that our action involves no more than first derivatives, we have, by the chain rule,

$$\delta\mathscr{L} = \frac{\partial\mathscr{L}}{\partial\phi^a}\delta\phi^a + \frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}\delta(\partial_\mu\phi^a) + \frac{\partial\mathscr{L}}{\partial x^\mu}\delta x^\mu . \tag{7}$$

Note that, for *functional* derivatives, we have the handy identity

$$\delta(\partial_\mu\phi^a) = \partial_\mu(\delta\phi^a) . \tag{8}$$

Therefore, counting the variation of both the measure and the integrand, we have

$$\delta S = \int d^D x \left\{(\partial_\mu\delta x^\mu)\mathscr{L} + \frac{\partial\mathscr{L}}{\partial\phi^a}\delta\phi^a + \frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}\partial_\mu(\delta\phi^a) + \partial_\mu\mathscr{L}\delta x^\mu .\right\} \tag{9}$$

Grouping together the first and fourth terms, and pulling the derivative in the third term past the canonical momenta $\Pi^\mu_a = \partial\mathscr{L}/\partial\partial_\mu\phi^a$ gives

$$\delta S = \int d^D x \left\{\partial_\mu\left[\mathscr{L}\delta x^\mu + \frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}\delta\phi^a\right] + \left[\frac{\partial\mathscr{L}}{\partial\phi^a} - \partial_\mu\left(\frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}\right)\right]\delta\phi^a\right\} . \tag{10}$$

Using the Euler-Lagrange equations to kill the second $[\cdots]$ term, and adding and subtracting a term $\Pi^\mu_a(\partial_\lambda\phi^a)\delta x^\lambda$, gives

$$\delta S = \int d^D x\, \partial_\mu \left\{\frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}\left[\delta\phi^a + (\partial_\lambda\phi^a)\delta x^\lambda\right] - \left[\frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}(\partial_\lambda\phi^a) - \delta^\mu_\lambda\mathscr{L}\right]\delta x^\lambda\right\} . \tag{11}$$

Defining a new quantity known as the **canonical energy-momentum tensor**

$$T^\mu{}_\lambda \equiv \frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}(\partial_\lambda\phi^a) - \delta^\mu_\lambda\mathscr{L} , \tag{12}$$

gives

$$\delta S = \int d^D x\, \partial_\mu \left\{\frac{\partial\mathscr{L}}{\partial\partial_\mu\phi^a}\Delta\phi^a - T^\mu{}_\lambda\delta x^\lambda\right\} \tag{13}$$

(For the case of curved spacetime, we would need to recruit the more powerful definition

$$T^{\rm GR}_{\mu\nu} = \frac{2}{\sqrt{-g}} \frac{\delta S}{\delta g^{\mu\nu}} \,, \tag{14}$$

and to use the formalism of Killing tensors to express conservation laws.)

So far, we have an OK-looking formula for the variation of the action. But in order to derive the sought-after conservation law, we need to know how the symmetry variations of the fields and of the spacetime coordinates are connected to the infinitesimal parameters of the continuous symmetry $\{\Delta\omega^A\}$. Note that a priori the parameter index label A has nothing to do with the field index label a; in particular, the number of values A can take is typically quite different from those for a. In this notation we can then write

$$\begin{aligned} \Delta x^\mu &\equiv \left(\frac{\Delta x^\mu}{\Delta\omega^A}\right) \Delta\omega^A \,, \\ \Delta\phi^a &\equiv \left(\frac{\Delta\phi^a}{\Delta\omega^A}\right) \Delta\omega^A \,. \end{aligned} \tag{15}$$

Accordingly, the functional variation of the action can be written

$$\delta S = \int d^D x \left[\partial_\mu \left\{ \frac{\partial \mathscr{L}}{\partial\partial_\mu\phi^a} \frac{\Delta\phi^a}{\Delta\omega^A} - T^\mu{}_\nu \frac{\Delta x^\nu}{\Delta\omega^A} \right\} \right] \Delta\omega^A \,. \tag{16}$$

Since this holds true for arbitrary parameters $\Delta\omega^A$, it follows that

$$\partial_\mu J^\mu{}_A = 0 \,, \tag{17}$$

where the conserved Noether current $J^\mu{}_A$ is defined as

$$J^\mu{}_A \equiv \frac{\partial \mathscr{L}}{\partial\partial_\mu\phi^a} \frac{\Delta\phi^a}{\Delta\omega^A} - T^\mu{}_\nu \frac{\Delta x^\nu}{\Delta\omega^A} \,. \tag{18}$$

You should check for yourself that using this continuity equation (17), Stokes' Theorem, and assuming that the spatial current falls off quickly enough at spatial infinity, gives a conserved charge Q_A:

$$\frac{dQ_A}{dt} = 0 \quad \text{where} \quad Q_A = \int d^d x J^0{}_A \,. \tag{19}$$

Here are some good exercises to try, to check your understanding. In the case of spacetime translations, you should find by working through all the steps explicitly that the conserved quantities are the spacetime momenta, because $T_{\mu\nu}$ is the flux of μ-momentum in the ν direction. For rotations and boosts, you should find that the corresponding conserved quantities are proportional to

$$Q^{\nu\lambda} = \int d^d x \left(T^{0\nu} x^\lambda - T^{0\lambda} x^\nu \right) \,, \tag{20}$$

i.e. first moments of the energy-momentum tensor. The space-space components Q^{ij} are indeed the familiar angular momenta. Check also that conservation of the space-time components Q^{0i} expresses the fact that the centre of mass moves at constant velocity.

Next, consider a free complex scalar field $\Phi = (\phi_1 + i\phi_2)/\sqrt{2}$, where $\phi_{1,2}$ are real scalar fields, and symmetry under $U(1)$ phase rotations $\Phi' = e^{iq\alpha}\Phi$. By writing out the variations under a phase rotation, you can show that the Noether current following from $U(1)$ symmetry is

$$J^\mu = -iq\left[\Phi^*(\partial_\mu \Phi) - (\partial_\mu \Phi^*)\Phi\right] . \tag{21}$$

Remember first quantization for scalar fields from QFT1? Using the standard expansion of the field operator in terms of plane waves with operator coefficients, you can show that the Noether charge is

$$Q = q\int d^d k(b_k^\dagger b_k - c_k^\dagger c_k)\,, \tag{22}$$

where b_k create particles and c_k create antiparticles with momentum k. It is also illustrative to work out the commutator of the charge operator Q with the fields Φ and $\Phi^\dagger$.

1.2 Lie groups and the Poincaré group

Our focus in this section is on continuous symmetries parametrized by a finite number of continuous parameters. Mathematically, a bunch of symmetry transformations forms a **group** if the following axioms are obeyed: (i) closure under group multiplication, (ii) associativity of group multiplication, (iii) existence of an identity, and (iv) existence of inverses for every group element. When group transformations act on fields in a well-defined fashion, the fields are said to carry a **representation** of the group G. Mathematically, a representation is a linear group action of G on a vector space V by invertible transformations $v \mapsto D(g)v$. It must obey linearity, $D(g)(\alpha v_1 + \beta v_2) = \alpha D(g)v_1 + \beta D(g)v_2$, and also the product rule, $D(g_1 g_2) = D(g_1)D(g_2)$; the rule for inverses is $D(g^{-1}) = D(g)^{-1}$. A representation is said to be faithful if $D(g)$ is the identity only for the identity element. Figuring out what kinds of representations are possible for a given symmetry group is known as tensor analysis.

The simplest example of group theoretic organization of symmetry information that we have already met in undergraduate quantum mechanics is the rotation group describing angular momenta. The spherical harmonics $|j,m\rangle$ carry a representation of the rotation group and obey the eigenvalue equations[2] $\vec{J}^2|j,m\rangle = j(j+1)|j,m\rangle$ and $J^3|j,m\rangle = m|j,m\rangle$, where j is the principal quantum number and m is the projection along the z-axis, with $|m| \leq j$. The ladder operators $J_\pm \equiv J^1 \pm iJ^2$ raise/lower the m quantum number, $J^\pm|j,m\rangle = |j,m\pm 1\rangle$ for $\pm m \neq j$ or 0 otherwise. Note that in this story, a role of primary importance is played by the square of the angular momentum $\vec{J}^2$, which commutes with each J^i. The algebra obeyed by the J^i is

$$[J^i, J^j] = i\epsilon^{ijk}J^k\,, i = 1,2,3. \tag{23}$$

The spin j is quantized to take integer or half-integer values.

A general Lie group transformation $U \in G$ in any representation can be written as

$$U = \exp\left(i\omega^A T^A\right)\,, \tag{24}$$

where the $\{T^A\}$ are known as the **generators** which live in the **Lie algebra** $\mathscr{G}$ of the Lie group G of transformations. The generators T^A are as numerous as the additive parameters

[2] We have suppressed factors of c and $\hbar$.

ω^A: the index A runs from 1 to $d(G)$, the **rank** of the symmetry group. From a mathematics perspective, the existence of this exponential map described in eq.(24) is nontrivial; as physicists, we will simply use it. It is important to note that the generators of symmetries T^A will generically act *differently* on fields than on coordinates. For example, a purely internal symmetry will only act nontrivially on the fields and not at all on the coordinates.

Regardless of the representation they are acting upon, the generators obey a Lie algebra

$$\left[T^A, T^B\right] = i f^{ABC} T^C \,, \tag{25}$$

where the f^{ABC} are known as the **structure constants** and are group-specific. Structure constants matter because everything important we need to know about a Lie group is encoded in its structure constants. The generators of a Lie algebra obey an identity known as the Jacobi identity,

$$[T^A, [T^B, T^C]] + [T^B, [T^C, T^A]] + [T^C, [T^A, T^B]] = 0 \,. \tag{26}$$

The **Cartan subalgebra** of a Lie algebra is the collection of generators that commute with all other generators.

There is a great variety of representations available for Lie groups of interest to physicists. The simplest kind of faithful representation is the **fundamental representation**, for which the fields ϕ transform like vectors under the action of group elements U behaving like matrices,

$$\phi' = U\,\phi \,. \tag{27}$$

These matrices are $d(f) \times d(f)$ matrices, where $d(f)$ is the dimension of the fundamental representation f. For many physics applications, such as the $SU(3)_c \times SU(2)_I \times U(1)_Y$ of the Standard Model of particle physics, quarks and leptons live in the fundamental representation and it is complex. As we can see by Taylor expanding (24), the infinitesimal form of a group transformation acting on a fundamental representation is

$$\Delta\phi = i\Delta\omega^A \left(T_f^A\right) \phi \,. \tag{28}$$

The above formula is handy when figuring out Noether currents, because it quickly gives us one of the puzzle pieces:

$$\frac{\Delta\phi^a}{\Delta\omega^A} = i \left(T_f^A\right)^a{}_b \, \phi^b \,. \tag{29}$$

For applications to gauge theories, another useful representation is the **adjoint representation**, which you will play with later in a homework assignment. For a field A living in the adjoint representation,

$$A' = U\,A\,U^{-1} \,, \tag{30}$$

from which we can quickly infer the infinitesimal form

$$\Delta A = i\Delta\omega^B [T_{Adj}^B, A] \,. \tag{31}$$

In the adjoint representation Adj, the generators are $d(G) \times d(G)$ matrices proportional to the structure constants, where $d(G)$ is the rank of the group,

$$(T_{Adj}^A)^{BC} = -i f^{ABC} \,. \tag{32}$$

The Lie algebra works out correctly in this adjoint representation because of the Jacobi identity. Since the structure constants are real and antisymmetric, $T^A_{Adj} = -(T^A_{Adj})^*$, and so the adjoint representation is real. Gauge potentials will turn out to live in the adjoint.

The normalization of the generators T^A is a matter of convention, and it is important to set a consistent convention because the Lie algebra eq.(25) is not invariant under a change of basis. In any representation, we can write

$$\mathrm{Tr}\left(T^A_r T^B_r\right) = \mathscr{C}(r)\delta^{AB}\,, \tag{33}$$

where $\mathscr{C}(r)$ is fixed for any given representation r and is known as the **index** of the representation. We can readily prove from this that

$$f^{ABC} = -\frac{i}{\mathscr{C}(r)}\mathrm{Tr}\left([T^A_r, T^B_r]T^C_r\right)\,. \tag{34}$$

Lie groups have an analogue of $\vec{J}^2$ for angular momentum, known as the **quadratic Casimir**, which can be readily shown to commute with all the T^A. It is built out of the sum of squares of the generators,

$$\sum_{A=1}^{d(G)} T^A_r T^A_r = C_2(r)\mathbb{1}_r\,. \tag{35}$$

In eq.(35), the matrix $\mathbb{1}_r$ is a $d(r)\times d(r)$ matrix, where $d(r)$ is the dimension of the representation r, and it must appear on the RHS of the above equation by Schur's lemma. Taking the trace of this formula yields the handy identity

$$d(G)\mathscr{C}(r) = d(r)C_2(r)\,. \tag{36}$$

For the $SU(N)$ groups which are heavily used in theoretical particle physics applications, the tradition is to choose $\mathscr{C}(f) = 1/2$,

$$\mathrm{Tr}\left(T^A_f T^B_f\right)\big|_{SU(N)} = \frac{1}{2}\delta^{AB}\,. \tag{37}$$

This convention is the one we will use in homework assignments and the final exam.

What kinds of Lie groups might we encounter[3] in theoretical physics? Matrix Lie groups include

- $GL(n)$, general linear: invertible
- $SL(n)$, special linear: $\det(S) = +1$
- $U(n)$, unitary: $U^\dagger = U^{-1}$
- $O(n)$, orthogonal: $O^T = O^{-1}$
- $SU(n)$, special unitary
- $SO(n)$, special orthogonal
- $Sp(n)$, symplectic: $SgS^\dagger = g$ where g is a $2n\times 2n$ matrix $g = \begin{pmatrix} 0 & +1 \\ -1 & 0 \end{pmatrix}$.

[3]Discrete groups and their irreps are pertinent to crystallography and other physics applications; we will not develop their representation theory here.

Note that the above notational conventions are specific to the Lie groups. The Lie *algebras* are normally denoted by the same symbols but with small letters rather than capital letters. For instance, the Lie algebra for $SU(3)$ is denoted $su(3)$.

A Lie group manifold that is finite-dimensional and compact is called a **compact** group. If the group has no $U(1)$ factors, it is called **semi-simple**. If in addition the algebra cannot be divided into two mutually commuting sets of generators, then it is **simple**. Back in the 19th century, Killing and Cartan classified all the compact simple Lie algebras, and they are known as the **classical groups**. They are: $SU(N)$, $SO(N)$, and $Sp(N)$ for arbitrary N, and the exceptional groups G_2, F_4, E_6, E_7, and E_8. We will not say anything more about the five exceptional groups except to mention that they are used in building GUTs and appear in compactifications of superstring theory. $SU(N)$ is all we need to know about to handle the Standard Model.

From the definition of the special orthogonal group, it is clear that matrices $M \in SO(D)$ obey

$$M^T \mathbb{1} M = \mathbb{1} \,. \tag{38}$$

When we are working in Lorentzian signature spacetime, we are interested in coordinate transformations that preserve the Minkowski norm rather than the Euclidean norm,

$$M^T \eta M = \eta \,, \tag{39}$$

where $\eta = \text{diag}(1, -1, \ldots, -1)$. Such matrices M are said to belong to the group $SO(1, d)$; in our 4D case, this is $SO(1, 3)$. More generally, if we preserved a Lorentzian norm with p $(+1)$ entries and q (-1) entries, we would have the group $SO(p, q)$. These groups are non-compact.

How about a basic example of a Lie group and a fundamental representation? Physicists always appreciate seeing a minimal working example of some concept. Consider first a simple $U(1)$ phase rotation of a complex field ϕ with parameter $\{\omega^A\} = \{\theta\}$. We know how this operates:

$$\phi' = e^{i\theta}\phi \,. \tag{40}$$

So in this case, the generator is just 1. That was easy! Let us also inspect a slightly less nontrivial example involving an external symmetry: translations in the x^1 direction. For this case, we also have only one parameter: $\{\omega^A\} = \{x^1\}$. For the 1D translation group, which is Abelian, the simplest nontrivial irreducible representation (or **irrep** for short) is a 1D plane wave: $\phi = e^{ip_1 x^1}$. When acting on this plane wave, the momentum operator p_1 can be represented as $p_1 = -i\partial_1$. Why does this work? First, the sign. We have set our conventions to ensure that the standard position-momentum commutators work: $[x^1, p_1] = i\hbar$. Second, this representation of the momentum operator gives the correct eigenvalue equation when acting on the 1D plane wave: $(-i\partial_1)\phi = p_1\phi$. So in this case, the generator of x^1 translations acting on 1D plane waves $e^{ip_1 x^1}$ is $-i\partial_1$, and the momentum eigenvalue is p_1. Note: if instead of a continuous translation symmetry we only had symmetry under translations by a lattice vector, we would get Bloch waves instead of plane waves.

A physically important property of the parameters of Lie groups is their **additivity**. Composition of two transformations along the same generator is described simply by adding their parameters. For rotations, this means we are talking about angles; for translations, the story is all about their vectors. For Lorentz boosts, however, velocity is *not* additive under composition. The correct (additive) parameter to use is the **rapidity** ζ, defined by

$$\frac{v}{c} = \tanh\zeta\,. \tag{41}$$

As you should check for yourself explicitly, rapidity adds simply under composition of Lorentz boosts. Note that boost rapidities obey $\zeta \in (-\infty, +\infty)$, a non-compact interval. Accordingly, the Poincaré group is a **non-compact** Lie group. By contrast, the rotation group is **compact**, as its angle parameters obey $\theta \in [0, 2\pi]$.

What is the group of symmetries symmetries of flat Minkowski spacetime? Consider our old friends position X^μ and momentum P^ν. As we saw above with the simple 1D plane wave example, the momentum can be thought of as the generator of translations,

$$P_\mu = -i\partial_\mu\,, \tag{42}$$

where the index parametrizes the directionality of the translation in spacetime. The total angular momentum generator $M_{\mu\nu}$ is defined[4] by

$$M_{\mu\nu} = L_{\mu\nu} + \Sigma_{\mu\nu} = (-X_\mu P_\nu + X_\nu P_\mu) + \Sigma_{\mu\nu}\,,$$

where $L_{\mu\nu}$ encodes the orbital angular momentum and $\Sigma_{\mu\nu}$ the spin angular momentum. The M_{0i} correspond to generators of Lorentz boosts, with the index i encoding the spatial directionality of the boost vector. The M_{ij} correspond to angular momentum generators: notice that each is defined by an antisymmetrized pair $[ij]$ corresponding to the plane of rotation for that angular generator.

The commutation relations for the generators of the Poincaré algebra are

$$\begin{aligned} [P_\mu, P_\nu] &= 0\,, \\ [P_\mu, M_{\rho\sigma}] &= +i\left(\eta_{\mu\rho}P_\sigma - \eta_{\mu\sigma}P_\rho\right)\,, \\ [M_{\mu\nu}, M_{\rho\sigma}] &= +i\left(\eta_{\nu\rho}M_{\mu\sigma} - \eta_{\mu\rho}M_{\nu\sigma} + \eta_{\mu\sigma}M_{\nu\rho} - \eta_{\nu\sigma}M_{\mu\rho}\right)\,. \end{aligned} \tag{43}$$

The Poincaré group, the symmetry group of flat $D = d + 1$ dimensional spacetime, is mathematically referred to as the **semi-direct product** of the Lorentz group $SO(1, d)$ with the translation group. It is a semi-direct product, rather than a direct product, because translation vectors get Lorentz-transformed under boosts and rotations: they behave, exactly as they should, like one-index tensors.

How do the $\Sigma^{\mu\nu}$ act? For scalar fields the action of spin generators $\Sigma^{\mu\nu}$ on them is, of course, trivial (they have no spin, so nothing happens). For the case of spin-half Dirac fermion representations, the spin generators are proportional to the two-index antisymmetrized product of gamma-matrices: $(\Sigma^{\mu\nu})_{\alpha\beta} = i([\gamma^\mu, \gamma^\nu])_{\alpha\beta}/4$. For spin one representations, the spin generators are also constant matrices, built of antisymmetric combinations of Minkowski metric tensors: $(\Sigma^{\rho\sigma})_{\mu\nu} = i(\eta^\rho{}_\mu\eta^\sigma{}_\nu - \eta^\sigma{}_\mu\eta^\rho{}_\nu)$.

[4] Our signature convention is mostly minus.

1.3 Origin of wave equations

Our story so far is independent of the dimensionality of spacetime and works for any spin. For the case of three spatial dimensions (and only for $d = 3$), we can write down an angular momentum pseudovector J_i defined by $J_i = \frac{1}{2}\epsilon_{ijk}M_{jk}$. (Note: these J_i commute with the Hamiltonian.) The boosts generators, by contrast, are defined by $M_{0i} = K_i$. (Note: the K_j do not commute with the Hamiltonian!) They have nontrivial commutators with the angular momenta:

$$[J_i, J_j] = +i\epsilon_{ijk}J_k\,, \qquad [J_i, K_j] = +i\epsilon_{ijk}K_k\,, \qquad [K_i, K_j] = -i\epsilon_{ijk}J_k\,. \tag{44}$$

Defining $N_i = (J_i + iK_i)/2$ and naïvely defining $N_i^\dagger = (J_i - iK_i)/2$ makes the algebra split apart:

$$[N_i, N_j] = i\epsilon_{ijk}N_k\,, \qquad \left[N_i^\dagger, N_j^\dagger\right] = i\epsilon_{ijk}N_k^\dagger\,, \qquad \left[N_i, N_j^\dagger\right] = 0\,. \tag{45}$$

This is a direct consequence of the mathematical fact that there is an isomorphism between $SO(4)$ and a direct product of two $SU(2)$s: $SO(4) \simeq SU(2) \times SU(2)$. Strictly speaking, it is only in Euclidean space that we have this split of the Lorentz group into a $SU(2) \times SU(2)$. The subtlety in Minkowski spacetime is that the boost generators K_i of the Lorentz algebra $so(1,3)$ are actually not Hermitean, because boosts involve a non-compact parameter. (The rotation generators are, meanwhile, Hermitean as their parameters are compact angles $\theta \in [0, 2\pi]$.) Why does this happen? Mathematically, unitary irreducible representations of non-compact groups are infinite-dimensional, and so to get finite-dimensional representations pertinent to physics we have to give up unitarity of boost matrices. Morally, we can think of the Minkowski case with Lorentz symmetry $SO(1,3)$ as a simple Wick rotation of the Euclidean case. This principle is sufficient to understand how classical and quantum fields transform under spacetime symmetries.

We can then characterize any irreducible representation of the Lorentz group by its angular momentum quantum numbers under each $SU(2)$. For a state $|j_L, j_R\rangle$, the total spin of the representation is $j_L + j_R$. Since both $j_{L,R}$ must be non-negative integers or half-odd-integers, the total spin is also of this character. This follows from the familiar rules for adding angular momenta, which themselves follow from the algebra of the J_i. For example, a scalar field is the $|0,0\rangle$ representation. A Weyl fermion is the $|\frac{1}{2}, 0\rangle$ or $|0, \frac{1}{2}\rangle$ (depending on handedness), while a Dirac fermion is the direct sum $|\frac{1}{2}, 0\rangle \oplus |0, \frac{1}{2}\rangle$; parity flips the two components. For spin one, our friend the vector potential A_μ corresponds to the $|\frac{1}{2}, \frac{1}{2}\rangle$, with the $\gamma^\mu_{\alpha\beta}$ playing the role of Clebsch-Gordan coefficient. The $|0,1\rangle$ and $|1,0\rangle$ correspond to the self-dual and anti-self-dual components of the field strength tensor $F_{\mu\nu}$ while the whole banana would be $|1,0\rangle \oplus |0,1\rangle$. The graviton is the $|1,1\rangle$ state. And so forth.

What is a wave equation for a field? It arises from analyzing how the field behaves under Poincaré transformations: combinations of translations, rotations, and boosts. More detail than we require may be found in the magnificent 3-volume "The Quantum Theory of Fields" text by Steven Weinberg, among other places. My exposition here is closely related to the simpler, more pedagogical exposition of Pierre Ramond in his classic text "Field Theory: A Modern Primer" on functional quantization methods. For now, I will just sketch quickly how it works for spin-half massless fermions.

How do spin-half massless fermions transform under Lorentz transformations? Recall that if we Wick rotate to Euclidean space, the 4D Lorentz group $SO(1,3)$ becomes $SO(4)$, which is isomorphic to $SU(2)\times SU(2)$. A nice basis for the generators of $SU(2)$ acting on spin-half fields is $\frac{1}{2}\sigma^i$, where the σ^i are the Pauli sigma matrices which obey

$$\sigma^i\sigma^j = \delta^{ij} + i\epsilon^{ijk}\sigma^k\,. \tag{46}$$

They are

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,, \quad \sigma^2 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}\,, \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,. \tag{47}$$

Then, Wick rotating back to Minkowski signature gives the following transformation laws. Left-handed spinors transform as

$$\psi_L(x) \to \Lambda_L\psi_L(x)\,, \tag{48}$$

where

$$\Lambda_L = \exp(i\vec{\sigma}\cdot[\vec{\theta} - i\vec{\zeta}]/2)\,. \tag{49}$$

In this expression, the $\vec{\theta}$ are the rotation parameters while $\vec{\zeta}$ are the boost parameters. Right-handers transform as

$$\psi_R(x) \to \psi_R'(x') = \Lambda_R\psi_R(x)\,, \tag{50}$$

where

$$\Lambda_R = \exp(i\vec{\sigma}\cdot[\vec{\theta} + i\vec{\zeta}]/2)\,. \tag{51}$$

A little algebraic fortitude yields an interesting fact: that under Lorentz transformations, $\sigma^2\psi_L^*$ transforms like a right-handed Weyl spinor. (To get the algebra to work out, it helps to show that $\Lambda_L^{-1} = \Lambda_R^\dagger$, $\sigma^2\Lambda_L\sigma^2 = \Lambda_R^*$, $\Lambda_L^T\sigma^2\Lambda_L = \sigma^2$, and similarly for $L \leftrightarrow R$.) More pertinently to our goal at hand, we can also show that two available fermion bilinears, $\psi_L^\dagger\psi_L$ and $\psi^\dagger\sigma^i\psi_L$, *mix* under boosts:

$$\psi_L^\dagger\psi_L \to \psi_L^\dagger\psi_L + \Delta\zeta^i\psi_L^\dagger\sigma^i\psi_L\,, \qquad \psi_L^\dagger\sigma^i\psi_L \to \psi_L^\dagger\sigma^i\psi_L + \Delta\zeta^i\psi_L^\dagger\psi_L\,. \tag{52}$$

Comparing the above with the general Lorentz transformation rule for 4-vectors

$$\Delta V^\mu = \Delta\epsilon^\mu{}_\nu V^\nu \tag{53}$$

where $\epsilon^{0i} = -\zeta^i$, we find that

$$(i\psi_L^\dagger\sigma^\mu\psi_L) = (i\psi_L^\dagger\psi_L, i\psi_L^\dagger\vec{\sigma}\psi_L) \tag{54}$$

is a bona fide four-vector under boosts. Showing that it is a bona fide four-vector under rotations is similarly straightforward. We can also show that

$$(i\psi_R^\dagger\bar{\sigma}^\mu\psi_R) = (i\psi_R^\dagger\psi_R, -i\psi_R^\dagger\vec{\sigma}\psi_R) \tag{55}$$

is a four-vector under boosts and rotations.

Then, using the above information, it is simple to build a Poincaré invariant action for a left- or right-handed Weyl fermion, by contracting with another handy four-vector lying around: ∂_μ. More precisely, in order to make an action for left-handers that is real, we need

$$S[\psi_L] = \int d^4x \, \frac{1}{2} \psi_L^\dagger \sigma^\mu \overleftrightarrow{\partial_\mu} \psi_L \,, \tag{56}$$

where

$$a \overleftrightarrow{\partial_\mu} b \equiv a(\partial_\mu b) - (\partial_\mu a) b \,. \tag{57}$$

Using this action, it is straightforward to write down the wave equation that follows for ψ_L. The same sort of logic works for ψ_R with $\sigma^\mu \to \bar{\sigma}^\mu$. Comparing the left- and right-handed wave equations to the massless limit of the Dirac equation familiar from PHY2403F, you will find that they are identical.

In summary, free particle wave equations arise from pure group theory – plus the assumption that our action is built from lowest nontrivial order in derivatives, in the spirit of effective field theory.

1.4 Origin of spin angular momentum and helicity

Casimir operators are symmetry generators which commute with all other symmetries. Let us now figure out the analogues of $\vec{J}^2$ for the Poincaré group. There will turn out to be two of them. To find the Casimirs, we need to identify those generators (which could be composites) which commute with everybody. An obvious candidate is the square of the momentum vector, which constitutes our first quadratic Casimir:

$$C_1 = P^\mu P_\mu \tag{58}$$

Since translations form an Abelian group, the Casimir nature of C_1 is guaranteed by the commutation relations for the Lorentz algebra. Of course, C_1 acts on states with momentum p^μ as[5]

$$P^\mu P_\mu |p^\mu\rangle = m^2 |p^\mu\rangle \tag{59}$$

It turns out to be a tad harder to find the second independent quadratic Casimir. To lay the foundations, consider the **Pauli-Lubański pseudovector** W^μ defined by

$$W^\mu \equiv \tfrac{1}{2} \epsilon^{\mu\nu\lambda\sigma} P_\nu M_{\lambda\sigma} \tag{60}$$

Notice that, by symmetry,

$$W^\mu P_\mu = 0 \tag{61}$$

and

$$W^\mu = \tfrac{1}{2} \epsilon^{\mu\nu\lambda\sigma} P_\nu \Sigma_{\lambda\sigma} \tag{62}$$

As you should check explicitly using the Poincaré algebra commutation relations, the second quadratic Casimir that commutes with everybody is

$$C_2 = W^\mu W_\mu \tag{63}$$

[5] We have suppressed the factors of c and $\hbar$.

As can be checked, these are the *only* two quadratic Casimirs for the Poincaré group. The remaining interesting question then becomes: what is the value of $W^\mu W_\mu$ evaluated on physical states? The answer to this turns out to depend sensitively on the value of m^2. There are four types of classes of momenta, originally classified by Eugene Wigner.

- $p^\mu p_\mu = m^2 > 0$, with $p^0 > 0$. These are positive-energy massive particles.
- $p^\mu p_\mu = m^2 = 0$, with $p^0 > 0$. These are positive-energy massless particles.
- $p^\mu = 0$. This corresponds to ... nothing. The vacuum.
- $p^\mu p_\mu = m^2 < 0$. These are either unphysical (free tachyons!) or virtual-only particles.

The **Little Group** for each class is defined to be the subgroup of Lorentz transformations leaving the momentum vector p^μ invariant. It characterizes the story of spin for a particle.

For the massive case, there exists a rest frame such that $\{p^\mu\} = \{m, \vec{0}\}$. Therefore, the little group for this case is $SO(d)$, the group of spatial rotations. Consider $d = 3$, the case with which you are intuitively familiar. The little group for this case is $SO(3)$, the familiar group of spatial rotations. Therefore, in a very deep sense, spin really is an angular momentum for massive particles[6]. To evaluate the value of the second quadratic Casimir here, consider again our Pauli-Lubański vector W^μ. We already know that $W^\mu P_\mu = 0$ by symmetry. Now make use of the form of the momentum in rest frame to find that $\{W^\mu\} = \{0, \vec{W}\}$. So, in the rest frame,

$$W_i = \tfrac{1}{2}\epsilon_{i0jk}P^0\Sigma^{jk} = -\tfrac{1}{2}m\epsilon_{0ijk}M^{jk} = \tfrac{1}{2}m\epsilon_{ijk}\Sigma^{jk} \equiv mc^2\Sigma_i \tag{64}$$

where Σ_i is a proper 3-vector. Therefore,

$$W^\nu W_\nu |p^\mu\rangle = m^2 s(s+1)|p^\mu\rangle \tag{65}$$

For the massless case, a representative momentum is $\{p^\mu\} = E\{1, 0, \ldots 0, 1\}$. Consider again $d = 3$ as a special case. For $d = 3$ the little group is $E(2)$, the Euclidean group of rotations in the (x, y) plane combined with x and y translations. How about the value of the quadratic Casimirs evaluated on a massless state? First, we already know that the momentum four-vector squares to zero. Second, we also know that the Pauli-Lubański pseudovector is orthogonal to the momentum vector. Third, because the mass is zero, the Pauli-Lubański pseudovector also squares to zero. As you should check for yourself, the only way for two null vectors to be orthogonal to one another in Lorentzian signature is for them to be parallel. Here this implies directly that

$$W_\mu = hP_\mu\,, \tag{66}$$

where h is a pseudoscalar known as the **helicity**. By the definition of W_μ and P_μ and the fact that h is a pseudoscalar, there can be only two allowed values for helicity: $h = +|s|$ and $h = -|s|$, where s is the spin. In this case, the spin data do not constitute an angular momentum at all, but are instead characterised by helicity alone.

[6]Mathematically, $SU(2)$ is isomorphic to the double cover of $SO(3)$; the double covering is why spin must be either integer or half-odd-integer.

2 Gauge Symmetry

2.1 Abelian gauge symmetry and QED

We know the free Dirac fermion action[7]

$$S_{1/2} = \int d^D x \left[i\bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi \right] \tag{67}$$

and the Maxwell action

$$S_1 = \int d^D x \left[-\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right] \tag{68}$$

Let us now try to work out the form of a Poincaré invariant interaction term between spin-half matter fermions and massless spin-one gauge fields. Our fermion fields are the spinor ψ and the conjugate spinor $\bar{\psi}$, which are in the $(0, \frac{1}{2}) \oplus (\frac{1}{2}, 0)$ and $(0, \bar{\frac{1}{2}}) \oplus (\bar{\frac{1}{2}}, 0)$ representations of $SO(1,3)$ respectively, while our vector is in the $(\frac{1}{2}, \frac{1}{2})$ representation. Therefore, we are in need of a Clebsch-Gordan coefficient that connects a spinor and a conjugate spinor to a vector. The solution to this problem is called the gamma-matrix. To lowest order in derivatives, the interaction lagrangian can be written as

$$S_{\rm int} = q \int d^D x \left(\bar{\psi} \gamma^\mu \psi A_\mu \right) \,, \tag{69}$$

where q is the electric charge. As you should check explicitly, another way to write the total action is

$$S_{\rm QED} = \int d^D x \left[\bar{\psi} \left\{ i \left(\mathbb{1}\partial\!\!\!/ - iqA\!\!\!/ \right) - \mathbb{1}m \right\} \psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right] \tag{70}$$

In these conventions, the gauge-covariant derivative on spinors is

$$D_\mu = \mathbb{1}\partial_\mu - iqA_\mu \,, \tag{71}$$

and so

$$S_{\rm QED} = \int d^D x \left[\bar{\psi} \left(iD\!\!\!\!/ - m \right) \psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \right] \tag{72}$$

A local phase transformations of matter fields

$$\psi \to e^{iq\alpha(x)} \psi(x) \tag{73}$$

can be compensated – locally – by gauge transformation of the connection

$$A_\mu \to A_\mu + \partial_\mu \alpha(x) \tag{74}$$

The idea of **gauge potential as compensator** turns out to be a very powerful one, reaching far beyond the case of $U(1)$ for QED. The key advantage of the gauge-covariant derivative is that it transforms **covariantly** under gauge transformations:

$$D_\mu \psi(x) \to e^{iq\alpha(x)} D_\mu \psi \tag{75}$$

[7] For the sticklers: this is equivalent to the previously written fermion action, upon integration by parts.

where

$$D_\mu = \partial_\mu - iqA_\mu \tag{76}$$

as above.

For the electron, the charge is $q = -e$. We then proceed to do perturbation theory as a series in e^2. More properly, reintroducing our units, this is a perturbation series in

$$\alpha_{\rm QED} \equiv \frac{e^2}{\hbar c} \tag{77}$$

2.2 Nonabelian gauge symmetry

Consider local gauge transformations on matter fields $\Psi(x)$ of the general form

$$\Psi \to \Psi' = U(x)\,\Psi \tag{78}$$

where $U(x)$ (or U for short) is a spacetime-dependent matrix and Ψ is in the fundamental representation. Generally such matrices U of the symmetry group will not commute with one another. Accordingly, this is called a non-Abelian gauge transformation.

Because of the spacetime dependence in the above gauge transformation, partial derivatives of matter fields do not transform covariantly. So let us introduce a compensator known as the **Yang-Mills gauge connection** which allows us to sensibly define parallel transport of matter fields. It is designed to ensure that the **gauge-covariant derivative** of a matter field transforms covariantly:

$$D_\mu\Psi \to (D_\mu\Psi)' = U(D_\mu\Psi)\,, \tag{79}$$

where we have defined

$$D_\mu\Psi = \partial_\mu\Psi - igA_\mu\Psi\,. \tag{80}$$

Here, the gauge field A_μ is an element of the Lie algebra:

$$A_\mu = A_\mu^A T^A\,. \tag{81}$$

More abstractly, the covariant derivative is

$$D_\mu = \mathbb{1}\partial_\mu - igA_\mu\,. \tag{82}$$

As can be straightforwardly checked, the required transformation law for the non-Abelian **gauge potential** A_μ is

$$A_\mu \to A'_\mu = UA_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1} \tag{83}$$

$$= UA_\mu U^{-1} + \frac{i}{g}U(\partial_\mu U^{-1})\,, \tag{84}$$

where in the last line we used the identity $\partial(UU^{-1}) = \partial(\mathbb{1}) = 0$. For infinitesimal parameters $\Delta\omega$, this transformation law reads

$$\Delta A_\mu = i\Delta\omega^C[T^C, A_\mu] + \frac{1}{g}T^C(\partial_\mu\Delta\omega^C)\,, \tag{85}$$

or in component form

$$\Delta A_\mu^C = \frac{1}{g}\left[\partial_\mu \Delta\omega^C + g f^{ABC} A_\mu^A \Delta\omega^B\right] . \tag{86}$$

How does the covariant derivative D_μ act on fields in the adjoint representation, such as the gauge potential and gauge field strength? Let us write a field in the adjoint as $\chi = \chi^C T^C$. Then

$$\begin{aligned}(D_\mu \chi)^C &= \left[(\mathbb{1}\partial_\mu - igA_\mu)\,\chi\right]^C \\ &= \left[\delta^{CA}\partial_\mu - igA_\mu^B (T^B)^{CA}\right]\chi^A \\ &= \partial_\mu \chi^C + g f^{ABC} A_\mu^A \chi^B ,\end{aligned} \tag{87}$$

because $(T^A)^{BC} = -if^{ABC}$. Therefore,

$$\Delta A_\mu^C = \frac{1}{g}(D_\mu \Delta\omega)^C . \tag{88}$$

An important property of gauge-covariant derivatives is that they obey the commutator

$$[D_\mu, D_\nu] = -igF_{\mu\nu} , \tag{89}$$

where the gauge field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu] . \tag{90}$$

Even though the gauge potential A_μ transforms like a connection under gauge transformations, as in (83), the field strength transforms covariantly,

$$F_{\mu\nu} \to F'_{\mu\nu} = U F_{\mu\nu} U^{-1} . \tag{91}$$

Note that F is not *in*variant under gauge transformations, but it is the closest thing possible in a non-Abelian gauge theory: *co*variant.

Note also that there are many other ways of introducing the Yang-Mills gauge potential and field strength. You can peruse QFT textbooks like the ones by Matthew Schwartz (modern; recommended), Michael Peskin and Daniel Schroeder, and Pierre Ramond to see some examples.

2.3 Yang-Mills Lagrangian and equations of motion

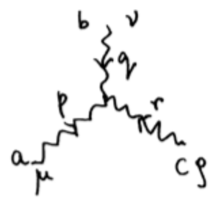
We can use the gauge field strength $F_{\mu\nu}$ to build ourselves a gauge-invariant Lagrangian, by tracing over the group indices[8]:

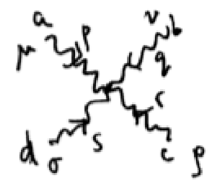
$$\mathscr{L}_{\rm YM} = -\frac{1}{2}\mathrm{Tr}(F^{\mu\nu}F_{\mu\nu}) = -\frac{1}{4}F^{\mu\nu\,A}F^A_{\mu\nu} . \tag{92}$$

Note three very important facts that make Yang-Mills qualitatively different than Maxwell. First, the gauge potential lives in the Lie algebra, $A_\mu = A_\mu^A T^A$. Second, the field strength

[8]The relative factor of 2 comes from the normalization convention (37) for $SU(N)$, $\mathrm{Tr}(T^A T^B) = \frac{1}{2}\delta^{AB}$.

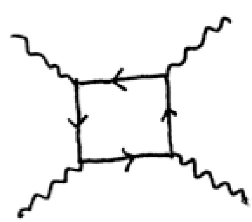
involves commutator terms quadratic in A, not just the covariant curl part linear in A like for Maxwell: $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]$. Third, the non-Abelian generators of the gauge symmetry T^A obey nontrivial commutation relations, $[T^A, T^B] = if^{ABC}T^C$. Putting these three facts together, we can see already that even the classical physics of Yang-Mills is very different from that for the Maxwell field: **the Yang-Mills Lagrangian contains self-interactions of both cubic and quartic type**.





This result is intimately related to the fact that superposition, a bedrock idea in electromagnetism, does not hold for Yang-Mills theories. It cannot: the field equations for A_μ^A are inherently nonlinear. Note that in momentum space a Feynman diagram must involve a power of momentum for the cubic self-coupling, by the definition of F in terms of A, whereas the quartic self-coupling involves no powers of momentum for the same reason.

By contrast, there are no classical cubic or quartic interaction terms for the photon at all. So the only way it is possible to get light-by-light scattering in QED is to involve quantum loops. For instance, a scattering amplitude involving (say) four external photon legs has to possess one internal loop with four fermion/antifermion propagators.



This diagram is $\mathscr{O}(\alpha_{EM}^2)$, which is higher-order than the classical four-point self-interaction for Yang-Mills gauge fields. This essential difference makes the Feynman graph expansion for non-Abelian gauge theory qualitatively different from that for QED. The nonlinearity turns out to be physically essential for describing the quantum dynamics of the strong and weak nuclear interactions.

Note: It is possible to perform a field redefinition

$$A_\mu = \frac{\tilde{A}_\mu}{g} \,. \tag{93}$$

In this convention, the factor of g disappears in the covariant derivative

$$D_\mu = \mathbb{1}\partial_\mu - i\tilde{A}_\mu \,, \tag{94}$$

and in the field strength

$$\tilde{F}_{\mu\nu} = \partial_\mu \tilde{A}_\nu - \partial_\nu \tilde{A}_\mu - i[\tilde{A}_\mu, \tilde{A}_\nu] \tag{95}$$

while the Yang-Mills Lagrangian develops the overall normalization out front

$$\mathcal{L}_{\rm YM} = -\frac{1}{2g^2}{\rm Tr}\left(\tilde{F}^{\mu\nu}\tilde{F}_{\mu\nu}\right) . \tag{96}$$

We will stick with the earlier convention, especially when coupling different types of leptons and quarks to the same $SU(3)_c \times SU(2)_I \times U(1)_Y$ Standard Model gauge fields.

Our action principle for Yang-Mills gauge theory has kinetic term

$$\mathscr{L}_{\rm YM} = -\frac{1}{2}{\rm Tr}(F^{\mu\nu}F_{\mu\nu}) = -\frac{1}{4}F^{\mu\nu}_A F^A_{\mu\nu} \tag{97}$$

where the components of the field strengths are obtained from $F_{\mu\nu} = F^A_{\mu\nu}T^A$,

$$F^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu + g f^{BCA} A^B_\mu A^C_\nu \,. \tag{98}$$

How would we couple in matter? Imagine fields like quarks and leptons in a fundamental (vector) representation of the gauge symmetry group. The gauge-invariant action is the one where we use the minimal coupling recipe: replace partial derivatives by gauge-covariant derivatives. The resulting Lagrangian for Yang-Mills coupled to matter is then

$$\mathscr{L}_{\rm YM} = i\bar\psi \not{D}\psi - m\bar\psi\psi - \frac{1}{2}{\rm Tr}(F^{\mu\nu}F_{\mu\nu}) \tag{99}$$

Note that the spinors in this expression have two kinds of indices: a spacetime spinor index α (which we almost always suppress) and an internal gauge index A. Just as naturally, the Yang-Mills fields have a spacetime index μ and an internal gauge index A.

The equations of motion following from the Yang-Mills action coupled to fundamental matter are, as you should check explicitly,

$$\left(\delta^{AC}\partial^\mu + g f^{BCA} A^{\mu B}\right) F^C_{\mu\nu} = -J^A_\nu \,, \tag{100}$$

where

$$J^{\mu A} = g\bar\psi T^A \psi \tag{101}$$

Or, in even shorter-hand notation,

$$D^\mu F_{\mu\nu} = -J_\nu \,. \tag{102}$$

The counterpart to the dynamical equation (100) for $A_\mu(x)$ is known as the **Bianchi identity**. This is a mathematical identity morally similar to the two current-free Maxwell equations. The Bianchi identity for Yang-Mills follows from the definition of F in terms of A,

$$\sum_{\rm cyclic} [D_\rho, [D_\mu, D_\nu]] = 0 \,, \tag{103}$$

or equivalently

$$D_\rho F_{\mu\nu} + D_\mu F_{\nu\rho} + D_\nu F_{\rho\mu} = 0 \,. \tag{104}$$

One final note. There is another term which we could have included in the Yang-Mills Lagrangian which is also quadratic in the field strength,

$$\mathscr{L}_* = \frac{\theta_*}{2}\epsilon^{\mu\nu\alpha\beta}\mathrm{Tr}\left(F_{\mu\nu}F_{\alpha\beta}\right) . \tag{105}$$

This can be shown to be a total derivative if the pseudoscalar θ_* is a constant,

$$\mathscr{L}_* = \theta_*\partial_\alpha\left[\epsilon^{\mu\nu\alpha\beta}\mathrm{Tr}\left(A_\beta\partial_\mu A_\nu - \frac{2ig}{3}A_\beta A_\mu A_\nu\right)\right] . \tag{106}$$

For now, we will ignore it. This is a great sin if we are interested in topological information encoded in the gauge field. Right at the end of the course when we talk about anomalies, we will reconnect to it briefly. If you are interested in learning about how gauge theory interfaces with mathematical topics like homotopy, anomalies, instantons, etc. there are many sources to learn about them. Some of the good introductory ones are Coleman's Aspects of Symmetry textbook and the classic review article by Eguchi, Gilkey and Hanson. Serious students of theoretical high-energy physics should study them.

2.4 The Standard Model, chirality, and gauging isospin and hypercharge

The gauge group of the standard model is $SU(3)_c \times SU(2)_I \times U(1)_Y$. The $SU(3)_c$ piece represents colour, the $SU(2)_I$ isospin, and the $U(1)_Y$ hypercharge. The electromagnetic $U(1)$ with which we are familiar is *not* one of these groups; instead, it will turn out to be a particular linear combination of the hypercharge and the diagonal component of isospin. All gauge bosons start life as massless vector bosons, as do the fermions in this model. It is the spontaneous symmetry breaking involving the scalar Higgs field which will give rise to both vector boson masses (via SSB of gauge symmetry) and fermion masses (via Yukawa couplings and SSB).

The Standard Model of particle physics is *chiral*. Left-handed and right-handed quarks and leptons couple differently, in particular to gauge bosons. Recall that the projectors

$$\mathscr{P}_\pm \equiv \frac{1}{2}(\mathbb{1} \pm \gamma_5) \tag{107}$$

project Dirac fermions onto their right- and left-handed components respectively.

Based on years of experimental results, it was eventually realized that left-handed leptons should be arranged as doublets of *weak isospin*

$$E_L = \begin{pmatrix} \nu_L \\ e_L \end{pmatrix} \tag{108}$$

Note that the fermion doublet has weak isospin $\frac{1}{2}$. The upper component of the doublet is taken to be the neutrino, which has $I^3 = +\frac{1}{2}$, while the charged lepton has $I^3 = -\frac{1}{2}$. Right-handed leptons, on the other hand, are $SU(2)$ singlets

$$(e_R) \tag{109}$$

Note that there is no right-handed neutrino (or left-handed antineutrino). Furthermore, as we will see shortly, neutrino masses are forbidden by $SU(2)_I \times U(1)_Y$ gauge invariance. This is why high-energy theorists refer to neutrino masses (and mixings) as "beyond the Standard Model" (sometimes abbreviated as BSM) physics.

What about quarks? They also have a similar structure, with

$$Q_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix} \tag{110}$$

while the right-handed quark fields are again singlets of $SU(2)$

$$(u_R) \quad (d_R) \tag{111}$$

Accordingly, the free fermion Lagrangian for the lightest generation may be written, with $SU(2)_I$ indices suppressed,

$$\begin{aligned} \mathscr{L} &= i\bar{e}_R\gamma\cdot\partial e_R + i\bar{e}_L\gamma\cdot\partial e_L + i\bar{\nu}_e\gamma\cdot\partial\nu_e + \\ &\quad + i\bar{u}_R\gamma\cdot\partial u_R + i\bar{u}_L\gamma\cdot\partial u_L + i\bar{d}_R\gamma\cdot\partial d_R + i\bar{d}_L\gamma\cdot\partial d_L \\ &= i\bar{e}_R\gamma\cdot\partial e_R + i\bar{E}_L\gamma\cdot\partial E_L + i\bar{u}_R\gamma\cdot\partial u_R + i\bar{d}_R\gamma\cdot\partial d_R + i\bar{Q}_L\gamma\cdot\partial Q_L \end{aligned} \tag{112}$$

Similarly for the muon and tau generations. Whatever form the hypercharge symmetry transformation takes, we know these terms would respect global $U(1)$ as each quark or lepton is matched with an antiquark or antilepton. The details of how it works with local gauge symmetry is our very next priority.

The above Lagrangian is invariant (so far) under a symmetry rotating the $SU(2)$ doublet around while leaving the singlet invariant. Promoting the $SU(2)_I$ to a gauge symmetry, we have

$$\begin{aligned} E_L &\overset{SU(2)_I}{\longrightarrow} e^{\frac{1}{2}i\sigma^i\alpha^i(x)}E_L \\ e_R &\overset{SU(2)_I}{\longrightarrow} e_R \end{aligned} \tag{113}$$

How are we going to find the $U(1)$ of electromagnetism when all we have so far in the way of $U(1)$s is hypercharge? Electric charge – for the left-handed fields at least – looks to be related to weak isospin and hypercharge via

$$Q = I^3 + Y \tag{114}$$

In other words, charge is basically the sum of the only two diagonal generators in the Lie algebra $su(2)\times u(1)$. Indeed, this will provide our definition of hypercharge in the following. In particular,

$$\begin{aligned} Q_\nu &= 0 = +\tfrac{1}{2} + Y(E_L) \\ Q_e &= -1 = -\tfrac{1}{2} + Y(E_L) \end{aligned} \tag{115}$$

which is consistent provided that

$$Y(E_L) = -\tfrac{1}{2} \tag{116}$$

What about the right-handers? For the singlet e_R, $I^3 = 0$ so that

$$Y(e_R) = -1 \tag{117}$$

We also need to know the hypercharge assignment for the quarks. We have

$$\begin{aligned} Q_u &= \tfrac{2}{3} = +\tfrac{1}{2} + Y(Q_L) \\ Q_d &= -\tfrac{1}{3} = -\tfrac{1}{2} + Y(Q_L) \end{aligned} \tag{118}$$

which is consistent provided that

$$Y(Q_L) = +\tfrac{1}{6} \tag{119}$$

How about the right-handed quarks? Since $I^3 = 0$, the right-handed up has hypercharge

$$Y(u_R) = +\tfrac{2}{3} \tag{120}$$

while the right-handed down has

$$Y(d_R) = -\tfrac{1}{3} \tag{121}$$

In summary, the $U(1)_Y$ hypercharge generator acts as

$$\begin{aligned} \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} &\overset{U(1)_Y}{\longrightarrow} \begin{pmatrix} e^{-\frac{1}{2}i\beta(x)} & 0 \\ 0 & e^{-\frac{1}{2}i\beta(x)} \end{pmatrix} \begin{pmatrix} \nu_e \\ e_L \end{pmatrix} \\ e_R &\overset{U(1)_Y}{\longrightarrow} e^{-i\beta(x)} e_R \\ \begin{pmatrix} u_L \\ d_L \end{pmatrix} &\overset{U(1)_Y}{\longrightarrow} \begin{pmatrix} e^{+\frac{1}{6}i\beta(x)} & 0 \\ 0 & e^{+\frac{1}{6}i\beta(x)} \end{pmatrix} \begin{pmatrix} u_L \\ d_L \end{pmatrix} \\ u_R &\overset{U(1)_Y}{\longrightarrow} e^{+\frac{2}{3}i\beta(x)} u_R \\ d_R &\overset{U(1)_Y}{\longrightarrow} e^{-\frac{1}{3}i\beta(x)} d_R \end{aligned} \tag{122}$$

Note that, unlike weak isospin, this transformation acts on both left- and right-handed fields, albeit with different charges (strengths of interaction).

In order to couple the gauge bosons of weak isospin and hypercharge to fermions, and to the Higgs boson, we will need to construct covariant derivatives. For $U(1)$, we introduce a *hypercharge gauge boson* X_μ and write gauge-covariant derivatives as

$$D_\mu^{(Y)} = \partial_\mu - ig' Y X_\mu \tag{123}$$

For the weak isospin part, we also define the *weak isospin gauge bosons*

$$D_\mu^{(I)} = \partial_\mu - ig \left(\frac{\sigma^i}{2}\right) W_\mu^i \tag{124}$$

These four gauge bosons, X_μ and W_μ^i, start life *massless.*

Putting together what we learned about Yang-Mills gauge theory and the above data on weak isospin and hypercharge for lepton and quark doublets and singlets, we obtain the non-Higgs part of the electroweak sector of the Standard Model for one generation:

$$\mathscr{L}_1 + \mathscr{L}_{1/2} = i\bar{e}_R\gamma^\mu \left(\partial_\mu + ig'X_\mu\right) e_R + i\bar{E}_L\gamma^\mu \left(\partial_\mu + \frac{1}{2}ig'X_\mu - \frac{1}{2}ig\sigma^i W_\mu^i\right) E_L$$

$$
\begin{aligned}
&+i\bar{u}_R\gamma^\mu\left(\partial_\mu - i\frac{2}{3}g'X_\mu\right)u_R + i\bar{d}_R\gamma^\mu\left(\partial_\mu + i\frac{1}{3}g'X_\mu\right)d_R \\
&+i\bar{Q}_L\gamma^\mu\left(\partial_\mu - \frac{1}{6}ig'X_\mu - \frac{1}{2}ig\sigma^i W^i_\mu\right)Q_L \\
&-\frac{1}{4}\left(\partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk}W^j_\mu W^k_\nu\right)^2 - \frac{1}{4}\left(\partial_\mu X_\nu - \partial_\nu X_\mu\right)^2
\end{aligned} \tag{125}
$$

The sad thing about this Lagrangian for describing the real world is that it gets all the masses wrong: they are all zero! Oops. Naively we might think that all is lost – putting in explicit mass terms for any of the desired particles would break gauge symmetry (ouch). This is where the magic of spontaneous symmetry breaking comes in. Shortly, we will see that by coupling in a versatile Higgs field transforming as a doublet under $SU(2)_I$ with a specific hypercharge under $U(1)_Y$, we will be able to generate masses for (a) the $W^\pm, Z$ vector bosons but *not* for the photon, and for (b) the leptons and quarks – without spoiling gauge symmetry of the model. The fact that we can get away with recruiting only one Higgs field to give all these particles masses is the group theoretic magic of the Glashow-Weinberg-Salam model that contributed to winning its eponymous physicists the Nobel Prize.

3 Spontaneous Symmetry Breaking

3.1 Goldstone's Theorem

Sometimes the vacuum of a QFT does not respect the symmetries of the action. This is called spontaneous symmetry breaking (SSB) and will turn out to be very important to a full understanding of the Standard Model of Particle Physics. SSB only occurs in systems with an infinite number of degrees of freedom. It is not possible in systems with a finite number of degrees of freedom.

A simple example that will be familiar is to take an iron bar and heat it up. A blob of molten Fe will have no overall magnetism. Now let the iron cool. The little Fe spins freeze into small aligned domains (and alignment can be encouraged with another magnet). The groundstate of the solid iron has all spins aligned, as it is a ferromagnet. When the iron cooled into a groundstate, it could have picked any direction to point the spins. It picked one randomly, and this broke rotational symmetry. Spontaneous symmetry breaking (SSB) like this could happen only because there was a continuous infinity of possible angles from which to choose. Another example would be a perfectly cylindrical pencil balanced vertically on its very tip. It has rotational symmetry about the axis, but waiting for gravity and random quantum fluctuations to do their thing will pick a direction – any direction, from 2π worth of angle.

We could label that continuous infinity of vacua $|\theta\rangle$ with angle(s) θ. Recasting the statement that the vacuum broke the rotational symmetry mathematically, we can say that $|\theta\rangle$ moves under the action of a rotation generator:

$$T_{\text{rot}}|\theta\rangle \neq 0 \tag{126}$$

SSB requires that the system under study be infinite. The reason is that, in order to know the angle(s) θ precisely, we would have to sum over all partial waves – a continuous infinity of them. It is the continuous infinity of possibilities that matters for spontaneous symmetry breaking.

Consider a scalar field theory with a symmetry and Lagrangian given by

$$\mathscr{L} = \tfrac{1}{2}(\partial\phi)^2 - V(\phi) \tag{127}$$

Suppose $\mathscr{L}$ is invariant under U-transformations. Then

$$\Delta V = \frac{\partial V}{\partial \phi^a}\Delta\phi^a \tag{128}$$

Therefore,

$$0 = \frac{\partial V}{\partial \phi^a}\left[i\Delta\omega^A (T^A)^a{}_b \phi^b\right] \tag{129}$$

Differentiating w.r.t. ϕ^c then gives

$$\frac{\partial^2 V}{\partial\phi^c \partial\phi^a}(T^A)^a{}_b\phi^b + \frac{\partial V}{\partial \phi^a}(T^A)^a{}_b\delta^b_c = 0 \tag{130}$$

This equation is the essence of what we are after.

Let us now evaluate this equation at a minimum of the potential, denoted ϕ_0 for convenience. Because of the structure of the action, derivatives are disfavoured: field gradients increase the energy. Similarly, $\partial V/\partial\phi^a(\phi{=}\phi_0){=}0$ on the equation of motion for the vacuum configuration. Therefore, evaluating (130) at a minimum gives

$$\langle\frac{\partial^2 V}{\partial\phi^c\partial\phi^a}\rangle(T^A)^a{}_b\langle\phi^b\rangle = 0 \tag{131}$$

where $\langle\ \rangle$ denotes the vacuum expectation value. Look at the first piece in this equation carefully. You should recognize it: it is none other than the mass matrix for fields $\{\phi^b\}$ at the minimum.

Let G denote the full symmetry group possessed by the action. Let H be the Little Group of $\langle\phi^b\rangle$: the symmetry transformations respected by the vacuum. Then G/H represents physical states. If $(T_A) \in H \subset G$, then $(T_A)^a{}_b\langle\phi^b\rangle = 0$ yielding an empty equation, because $\langle\phi^b\rangle$ respects H-symmetry. But if $T_A \notin H$, i.e. $(T_A)^a{}_b\langle\phi^b\rangle \neq 0$, then it is termed a "broken generator" [of symmetry] and by equation (130,131) there is a corresponding massless mode. In other words, we have just found

> **Goldstone's Theorem: in systems with spontaneous symmetry breaking, for every generator of the symmetry group broken by the vacuum there is a corresponding massless field called the "Nambu-Goldstone Boson".**

Physically, the most important thing about the Nambu-Goldstone Boson is that it is massless, i.e., it has a dispersion relation $E = |\vec{p}|$. An example of a Nambu-Goldstone boson is the phonon in fluids.

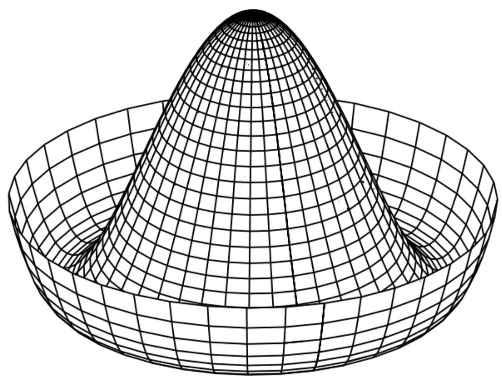
Technically, **Goldstone's Theorem** is already a powerful one at tree level. What is even more remarkable is that it **holds solid even when quantum corrections are taken into account**. This is shown in significant detail in Peskin and Schroeder's §11. The essence of why this works is the formulation we employ: functional quantization. Symmetries are kept manifest in the action, the generating functional, and the Green's functions of the theory. There is no need to break gauge symmetry, Lorentz symmetry, or whatever other symmetry may be present, in order to perform quantization. The nontrivial work amounts to showing that classical symmetries persist in the quantum action $\Gamma[\phi]$ when loop divergences are regularized consistently. Peskin and Schroeder §11 contains a very nice full discussion of this physics, starting with explicit loop calculations in the linear sigma model example in §11.2 and working up to a more abstract derivation in §11.6.

3.2 SSB with global symmetry

Consider a collection of real scalar fields $\{\phi^i\}$, where $i = 1, \cdots N$. Suppose that the potential for these fields has a global symmetry that makes it depend only on the magnitude of $\vec{\Phi}$, but not on the direction. For example:

$$S[\vec{\Phi}] = \int d^D x \left\{ \tfrac{1}{2}\partial^\mu\phi^i\partial_\mu\phi^i + \tfrac{1}{2}\mu^2\phi^i\phi^i - \lambda(\phi^i\phi^i)^2 \right\} \tag{132}$$

as illustrated in the picture below. Note the positive quartic and *negative* quadratic coefficients in $V(\vec{\Phi})$. This is not a typo; without the negative mass-squared term, SSB will not occur in this model.



An important physics point here is that the negative mass-squared is *not* a violation of causality – it simply indicates an instability, the endpoint of which *is* well-defined physically! The type of tachyon which does cause heart-stopping arrhythmias in theoretical high-energy physicists is the *free* tachyon, which just propagates and propagates, flouting causality all the while.

A shorthand for the above Lagrangian is

$$\mathscr{L} = \tfrac{1}{2}\partial^\mu \vec{\Phi} \cdot \partial_\mu \vec{\Phi} + \tfrac{1}{2}\mu^2 |\vec{\Phi}|^2 - \lambda |\vec{\Phi}|^4 \tag{133}$$

which is invariant under rotations of $\vec{\Phi}$. You should think of $\vec{\Phi}$ here as a vector in field space. The rotational symmetry acts as

$$\vec{\Phi} \to R\vec{\Phi}, \quad \text{i.e.} \quad \phi^i \to R^i{}_j \phi^j \tag{134}$$

Since $\phi^i \in \mathbb{R}$, the matrices R are orthogonal ($R^{-1} = R^T$).

How many generators are there? Our group transformations R and algebra generators T_A are connected via our friend the exponential map:

$$R = \exp\left(i\Delta\omega^A T_A\right) \tag{135}$$

In order for a rotation transformation R to be orthogonal, as you can quickly check, the generators of $o(N)$ must be antisymmetric:

$$(T_A)_{ij} = -(T_A)_{ji} \tag{136}$$

Counting the number of independent components of a real antisymmetric matrix, we get $\frac{1}{2}N(N-1)$. Naturally, this matches the number of rotations that we can make in N directions of field space.

The lowest-energy configuration for our linear sigma model is the one with

$$\vec{\Phi}(x^\mu) = \vec{\Phi}_0 \quad \text{and} \quad \left.\frac{\partial V}{\partial \vec{\Phi}}\right|_{\vec{\Phi}=\vec{\Phi}_0} = \vec{0} \tag{137}$$

Using the form of the action, we find quickly

$$(\vec{\Phi}_0)^2 = \frac{\mu^2}{4\lambda} \equiv v^2 \tag{138}$$

where v is short for VEV, vacuum expectation value. For definiteness, let us point the vev vector along the Nth direction in field space. Next let us perform a field redefinition:

$$\phi^i(x) = \left\{\pi^k(x), v + \sigma(x)\right\} \tag{139}$$

As you should check as an exercise, it follows straightforwardly that

$$\begin{aligned} \mathscr{L} &= \frac{1}{2}\partial^\mu \pi^k \partial_\mu \pi^k + \frac{1}{2}\partial^\mu \sigma \partial_\mu \sigma - \frac{1}{2}(2\mu^2)\sigma^2 \\ &\quad -2\sqrt{\lambda}\mu\sigma^3 - 2\sqrt{\lambda}\mu\pi^k\pi^k\sigma - \lambda\sigma^4 - 2\lambda\sigma^2\pi^k\pi^k - \lambda\pi^k\pi^k\pi^j\pi^j \end{aligned} \tag{140}$$

Notice that the π^k fields are massless; the only massive field left after spontaneous symmetry breaking is the σ field parametrizing field fluctuations over and above vacuum values. There are cubic and quartic self-interactions for the σ, a quartic self-interaction for the πs, and a cubic $\pi - \pi - \sigma$ coupling as well as a quartic $\pi - \pi - \sigma - \sigma$ coupling.

This exposition was for actions with "global" symmetry. In the case with gauge symmetry, the story has some extra twists to it, which are rather splendid physically. They also form the backbone of how the Higgs spontaneously breaks electroweak symmetry down to $SU(2) \times U(1)$ at low energies.

3.3 SSB with local gauge symmetry

Consider a single complex scalar field Φ with charge $(-e)$. Give it a wrong-sign mass term and a quartic self-interaction in its Lagrangian, as in our linear sigma model example. Now let us add a significant extra degree of difficulty by demanding not a global symmetry but a local "gauge symmetry". This requires that the Φ field be coupled to the A_μ field of electromagnetism. Local phase transformations acting on the scalar field

$$\Phi \to \Phi' = e^{-ie\Lambda(x)}\Phi \tag{141}$$

are compensated by a shift in the gauge connection

$$A_\mu \to A_\mu + \partial_\mu \Lambda\,. \tag{142}$$

Since you have met the $U(1)$ gauge theory of electromagnetism before, you already know that the simplest gauge-invariant action for Φ coupled to the gauge field is obtained via "minimal coupling":

$$\partial_\mu \to (\mathbb{1}\partial_\mu + ieA_\mu) \tag{143}$$

This is required to ensure gauge invariance. The resulting Lagrangian is

$$\mathscr{L} = ((\partial^\mu - ieA^\mu)\Phi^*)\left((\partial_\mu + ieA_\mu)\Phi\right) + \mu^2\Phi^*\Phi - \lambda(\Phi^*\Phi)^2 - \tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{144}$$

In analogy to how our previous example panned out, let us anticipate SSB and write

$$\Phi(x) = v + \frac{1}{\sqrt{2}} (\phi_1 + i\phi_2) \tag{145}$$

Physically, this states that the vev would be v and the two real fields $\phi_{1,2}$ parametrize fluctuations above the vacuum. In these variables,

$$\begin{aligned} \mathscr{L} &= -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + e^2v^2A^\mu A_\mu + \tfrac{1}{2}(\partial^\mu\phi_1)(\partial_\mu\phi_1) + \tfrac{1}{2}(\partial^\mu\phi_2)(\partial_\mu\phi_2) \\ &\quad -2\lambda v^2\phi_1^2 + \sqrt{2}eA^\mu\partial_\mu\phi_2 + \text{(cubic terms)} + \text{(quartic terms)} \end{aligned} \tag{146}$$

Something very important has just happened here: the photon got a mass! Note that the fluctuating field ϕ_1 has a mass term $-2\lambda v^2\phi_1^2$, while ϕ_2 has no mass term and is derivatively coupled. **Physically, we say that the photon has eaten the would-be Goldstone boson and become massive.** ϕ_2 does not have a direct physical interpretation as a field whose quanta we can measure in a detector. It is, in fact, possible to gauge away ϕ_2 entirely, as we shall now see explicitly.

Since mathematically $U(1) \simeq SO(2)$, our $U(1)$ gauge transformation of Φ amounts to an $SO(2)$ rotation in the (ϕ_1, ϕ_2) plane. Using (141) (142) and (145) we write the components of Φ in **unitary gauge** as a gauge-transformed vev in the ϕ_2 direction. Infinitesimally,

$$\begin{aligned} \phi_1' &= \phi_1 - e\Lambda\phi_2 \\ \phi_2' &= e\Lambda\phi_1 + \phi_2 + \sqrt{2}e\Lambda v \end{aligned} \tag{147}$$

(You can of course very quickly find the finite version of this equation too: you should get cosines and sines.) The critical observation is that we have (just) enough gauge freedom to use $\Lambda(x)$ to make $\phi_2(x)$ identically zero everywhere. You should perform the exercise to see this in very explicit terms yourself. Then, relabelling ϕ_1' to be just ϕ, we obtain

$$\mathscr{L} = -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} - e^2v^2A^\mu A_\mu + \tfrac{1}{2}(\partial^\mu\phi)(\partial_\mu\phi) - 2\lambda v^2\phi^2 + \text{(cubic terms)} + \text{(quartic terms)} \tag{148}$$

Our Lagrangian now contains only two fields: A_μ and ϕ. No trace of the eaten would-be Goldstone boson remains in unitary gauge. This was of course done totally on purpose.

Let us now briefly make contact with the Higgs mechanism in a condensed matter physics context. Consider a static configuration in our Abelian Higgs model above. The (negative of the) Lagrangian would then reduce to

$$-\mathscr{L}_{\text{static}} \equiv F = \tfrac{1}{2}\left|(\vec{\nabla} - ie\vec{A})\Phi\right|^2 + (-\mu^2)\left|\Phi\right|^2 + \lambda\left|\Phi^*\Phi\right|^2 + \tfrac{1}{4}\left|\vec{\nabla}\times\vec{A}\right|^2 \tag{149}$$

(Note: in our signature convention, ∂_μ and A^μ are decomposed as $\{\partial_\mu\} = \{\partial_0, \vec{\nabla}\}$ while $\{A^\mu\}=\{A_0, \vec{A}\}$ so that $\{A_\mu\}=\{A_0, -\vec{A}\}$.) This quantity F is exactly the same as the Landau-Ginzburg free energy in condensed matter physics, with the role of the wrong-sign mass parameter played by

$$\mu^2 = v(T_c - T) \tag{150}$$

near the critical temperature T_c and with the role of Φ played by the macroscopic many-particle wavefunction whose use is justified by BCS theory. When $T > T_c$, the minimum free energy is at $|\Phi| = 0$. By contrast, when $T < T_c$, the mass term becomes "tachyonic". We can analyze the physics that ensues by focusing on the current that is conserved because of $U(1)$ gauge symmetry,

$$J_\mu = ie\Phi^* \overleftrightarrow{\partial_\mu} \Phi - 2e^2 |\Phi|^2 A_\mu \, . \tag{151}$$

The spatial component of this is the usual 3-vector current $\vec{j}$. When $T < T_c$ and Φ varies only a teeny bit across the physical system, the second term in the current is strongly dominant. It gives

$$\vec{j} \simeq -\frac{e^2 \mu^2}{\lambda} \vec{A} \equiv -k^2 \vec{A} \, . \tag{152}$$

This is the London equation. It shows that the resistance must be zero, as $\vec{E} = \vec{j} R$ and $\vec{E} = -\partial \vec{A}/\partial t$. It is also easy to derive the Meissner effect, viz. explusion of magnetic flux. Starting from Ampère's Law $\vec{\nabla} \times \vec{B} = \vec{j}$ and using the Bianchi identity $\vec{\nabla} \cdot \vec{B} = 0$ gives $\nabla^2 \vec{B} = +k^2 \vec{B}$. So $\vec{B}$ has exponential falloff in position space, indicating that the magnetic field only penetrates to a characteristic depth set by $1/k$. Also, $\nabla^2 \vec{A} = +k^2 \vec{A}$, which would in relativistic form suggest a photon mass k. Again, this just the Higgs mechanism at work.

Our first way for symmetry of a complex scalar field theory to break spontaneously was for it to rest on *global* symmetry. In that case, we had

- Goldstone mode
- Lagrangian had 2 massive scalars
- SSB gave 1 massless scalar + 1 massive scalar

In the case of spontaneous symmetry breaking for a charged scalar field coupled to an Abelian *gauge* symmetry, we ended up with

- Higgs mode
- Lagrangian had 2 massive scalars and 1 massless gauge field
- SSB gave 1 massive gauge field + 1 massive scalar

Notice the remarkable thing that happened here – the photon ate the second scalar field and became massive! Accordingly, the second scalar plays the role of the longitudinal polarization of the gauge boson. This is nowadays called the (Abelian) Higgs phenomenon, although several important names should be credited as well, including Philip Anderson. The J.J. Sakurai Prize for Theoretical Particle Physics in 2010 was awarded to {Robert Brout & François Englert}, {Gerald Guralnik, Carl Hagen, & Tom Kibble}, and Peter Higgs "for elucidation of the properties of spontaneous symmetry breaking in four-dimensional relativistic gauge theory and of the mechanism for the consistent generation of vector boson masses". Two of them, Englert and Higgs, won the Nobel Prize in Physics 2013 "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider".

The gauge defined by equation (147) is known as the unitary gauge. It is very handy for keeping track of only the physical degrees of freedom and unitarity. On the other hand, unfortunately in non-Abelian gauge theories the Green's functions in unitary gauge do not have a renormalizable perturbation expansion – although the S-matrix does. It is therefore traditional to use the R_ξ gauges, which are well-suited to this task, defined by $G^A = \partial^\mu A^A_\mu + \xi g v \phi^A_2 = 0$, i.e.

$$\mathscr{L}_{\rm gf} = -\frac{1}{2\xi}G^A G^A = -\frac{1}{2\xi}\left(\partial^\mu A^A_\mu + \xi g v \phi^A_2\right)^2 \tag{153}$$

These gauges are a kind of interpolation between unitary gauge and a different class of gauge where renormalizability at loop level is easier to track. We will be able to explain this after we have introduced Fadeev-Popov ghosts.

We now turn to the case of a non-Abelian gauge symmetry, to illustrate the essential physics needed to explain electroweak symmetry breaking mechanism in the Standard Model. Consider the $O(N)$ model, with N real scalar fields ϕ^I, where $N = 3$. Then

$$\mathscr{L} = \tfrac{1}{2}D^\mu\phi^I D_\mu\phi^I + \tfrac{1}{2}\mu^2\phi^I\phi^I - \lambda(\phi^I\phi^I)^2 - \tfrac{1}{4}F^{\mu\nu I}F^I_{\mu\nu} \tag{154}$$

where

$$\begin{aligned} D_\mu\phi^I &= \partial_\mu\phi^I + g\epsilon^{JKI}A^J_\mu\phi^K \\ F^I_{\mu\nu} &= \partial_\mu A^I_\nu - \partial_\nu A^I_\mu + g\epsilon^{JKI}A^J_\mu A^K_\nu \end{aligned} \tag{155}$$

The reason why we are seeing ϵ^{IJK} appearing is that they are the structure constants of $O(3)$. These are exactly the same structure constants as for the group $SO(3)$, for a good reason. As you may wish to prove for yourself, orthogonal matrices O [over the reals] obey $\det(O) = \pm 1$. The part of the $O(3)$ group connected to the identity matrix is $SO(3)$; in $D = 3$, the rest of $O(3)$ is obtained via a parity transformation. Therefore, the structure constants for $o(3)$ are the same as for $so(3)$. Note also that, at the level of the Lie algebras, $su(2)$ and $so(3)$ are isomorphic. That is why we get ϵ^{IJK} for $so(3)$ as well as for $su(2)$. At the level of the Lie groups, $SU(2)$ and $SO(3)$ differ. In fact, $SU(2)$ is the double cover of $SO(3)$. Again, this is directly related to the fact that spin occurs for free fields only in integer or odd-half-integer values (in units of $\hbar$).

The quartic potential with wrong-sign mass term has a minimum at

$$|\phi_0| = \frac{\mu}{2\sqrt{\lambda}} \tag{156}$$

As with our Abelian case, let us choose the vacuum to lie along

$$\vec{\phi}_0 = v\hat{e}_3 \tag{157}$$

The physical fields are then: $\{\phi_1, \phi_2, \phi_3 - v \equiv \chi\}$. As you can check, straightforward algebra gives

$$\begin{aligned} \mathscr{L} &= \frac{1}{2}\left[(\partial^\mu\phi^1)(\partial_\mu\phi^1) + (\partial^\mu\phi^2)(\partial_\mu\phi^2) + (\partial^\mu\chi)(\partial_\mu\chi)\right] + vg\left[(\partial_\mu\phi_1)A^\mu_2 - (\partial_\mu\phi_2)A^\mu_1\right] \\ &\quad + \frac{1}{2}v^2g^2\left[(A^{1\mu}A^1_\mu) + (A^{2\mu}A^2_\mu)\right] - \frac{1}{4}(\partial_\mu A^I_\nu - \partial_\nu A^I_\mu)^2 \\ &\quad + 4v^2\lambda\chi^2 + (\text{cubic}) + (\text{quartic}) \end{aligned} \tag{158}$$

Notice what has happened here: only χ got a mass amongst the scalars, while $A_\mu^{1,2}$ have become massive. In other words, there are two would-be Nambu-Goldstone bosons which have been eaten.

In unitary gauge, we set

$$\vec{\phi}(x) = \hat{e}_3 \phi_3(x) = \hat{e}_3 (v + \chi) \tag{159}$$

everywhere. There is just enough local $O(3)$ gauge symmetry to accomplish this feat, at every x^μ. Then, in unitary gauge only,

$$\sum_{I=1}^{3} \left(D^\mu \phi^I\right)\left(D_\mu \phi^I\right)\Big|_U = v^2 g^2 \left(A_\mu^1 A^{1\mu} + A_\mu^2 A^{\mu 2}\right) + \left(\partial^\mu \chi\right)\left(\partial_\mu \chi\right) \tag{160}$$

and

$$\begin{aligned} \mathscr{L}|_U &= -\frac{1}{4}\left(\partial^\mu A^{\nu I} - \partial^\nu A^{\mu I}\right)\left(\partial_\mu A_\nu^I - \partial_\nu A_\mu^I\right) - \frac{1}{2} v^2 g^2 \left(A^{\mu 1} A_\mu^1 + A^{\mu 2} A_\mu^2\right) \\ &\quad + \frac{1}{2}\partial^\mu \chi \partial_\mu \chi - 4 v^2 \lambda \chi^2 + (\text{cubic}) + (\text{quartic}) \end{aligned} \tag{161}$$

We should now pause to check that we have a proper accounting for the degrees of freedom of this $O(3)$ linear sigma model in both the Higgs and Goldstone modes. Before any SSB, of course, we would have had 3 massive scalars and 3 massless vectors, giving $(3 \times 1) + (3 \times 2) = 9$ physical degrees of freedom. For SSB modes in the $O(3)$ linear sigma model, then,

Higgs

- 1 massive scalar (χ)
- 2 massive vectors (A_μ^1, A_μ^2)
- 1 massless vector (A_μ^3)
- Total: $(1 \times 1) + (2 \times 3) + (2) = 9$

Goldstone

- 1 massive scalar
- 2 massless scalars
- 3 massless vectors
- Total: $(1 \times 1) + (2 \times 1) + (3 \times 2) = 9$

The interesting thing is that this relatively innocuous looking $O(3)$ model possesses all the important features of a non-Abelian gauge field theory coupled to a a Higgs field, giving rise to SSB, Goldstone fields getting eaten and fattening their respective gauge bosons in a manner consistent with local gauge invariance. Here we even had one remaning unbroken generator, so that one scalar remained massive and the A_μ^3 component of the non-Abelian gauge field stayed massless. This feature is mission-critical, as the photon is known to be massless to an extremely high degree of experimental accuracy.

3.4 The Higgs boson

One of the main reasons for introducing the Higgs is to produce massive vector bosons in a manner consistent with gauge invariance. This is achieved via spontaneous breaking of the $SU(2)_I \times U(1)_Y$, leaving the electromagnetic gauge boson massless while giving masses to $W^\pm, Z$. Another equally important feature is that the Higgs gives mass to fermions in the Standard Model, quarks and leptons included.

For a fermion mass term, we need something that couples left- and right-handed fermion fields. A plain old ordinary mass term for the fermions would be forbidden by $SU(2)_I \times U(1)_Y$ gauge symmetry, so we need another way forward.

Let us first discuss the case of the leptons. Since E_L is an $SU(2)_I$ doublet while e_R is a singlet, we will need a [conjugate] $SU(2)_I$ doublet representation in order to have any hope of making an $SU(2)_I$-invariant Lagrangian from E_L, e_R . i.e.,

$$I(\Phi) = +\tfrac{1}{2}\,. \tag{162}$$

Indeed, a *Yukawa coupling* between the Higgs scalar and the electron-generation fermions of the form

$$\mathscr{L}_{\text{Yuk}} \supset -\lambda_e \left(\bar{E}_L \Phi e_R + \text{h.c.}\right) \tag{163}$$

would achieve the desired mass term structure – provided that Φ were to develop a vacuum expectation value through spontaneous symmetry breaking. This term respects $SU(2)_I$ as written. We also need to make sure that our Lagrangian for fermion-Higgs interactions is $U(1)_Y$ invariant. Left- and right-handed fermions differ in hypercharge: $Y(\bar{E}_L) = +1/2$, $Y(e_R) = -1$. Accordingly, the Higgs must transform in the following way to respect hypercharge gauge symmetry:

$$Y(\Phi) = +\frac{1}{2}\,. \tag{164}$$

Because $Q = I^3 + Y$, the upper component of the Higgs doublet will have charge $Q = +1$ and the lower component will have $Q = 0$.

Now we must check that these charge assignments also work for the quark sector while continuing to respect $SU(2)_I \times U(1)_Y$ gauge invariance! Let us see if we can write down a consistent Higgs-quark Yukawa coupling. As it happens, an accident of group theory for $SU(2)_I$ lets us get away with it. For the down-type quarks, we can write

$$\mathscr{L}_{\text{Yuk}} \supset -\lambda_d \bar{Q}_L \Phi d_R + \text{h.c.}\,. \tag{165}$$

In order for this coupling to respect $U(1)_Y$ as well as $SU(2)_I$, we need for the total hypercharge of this Yukawa term to vanish. We know that $Y(Q_L) = +1/6$ and that $Y(d_R) = -1/3$, so that indeed the Higgs with $Y(\Phi) = +1/2$ will do the trick. For the up-type quarks, we have to try something different to respect hypercharge and isospin invariance. This time, let us write

$$\mathscr{L}_{\text{Yuk}} \supset -\lambda_u \bar{Q}_L \tilde{\Phi} u_R + \text{h.c.}\,, \tag{166}$$

where the putative doublet field $\tilde{\Phi}$ must have the *opposite* hypercharge to Φ, because $Y(u_R) = +2/3$ and $Y(\bar{Q}_L) = -1/6$ and so $Y(\tilde{\Phi}) = -1/2$. The fortunate group theoretic accident of $SU(2)_I$ is that for an isospin-half doublet, we *can* make another isospin-half doublet with the opposite charge from the original simply by recruiting

$$\tilde{\Phi} = i\frac{\sigma^2}{2}\Phi^*\,. \tag{167}$$

You can check for yourself that this works or look it up in a group theory for physicists textbook. This trick does *not* work for other gauge groups. It is quite remarkable that we only had to use *one* Higgs field here, in order to write down Yukawa couplings that via

SSB would give masses to the electron-type lepton and both up-type and down-type quarks but not to the neutrino. In a more general theory, for example in SUSY QFTs, it will be necessary to have at least two Higgses, one up-type and one down-type.

In sum, the combined $SU(2)_I \times U(1)_Y$ invariance tells us to put the Higgs in an $SU(2)_I$ doublet; its upper component is the positively charged Higgs while its lower component is the neutral Higgs:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{168}$$

Since this is a complex scalar doublet, it has four real degrees of freedom.

Since we have already constructed gauge-covariant derivatives in general, and worked out the isospin and hypercharge assignments of the Higgs in particular, we can now write down a Lagrangian for the Higgs which exhibits SSB and the right kinds of gauge invariance:

$$\begin{aligned} \mathscr{L}_0 &= (D^\mu \Phi)^\dagger (D_\mu \Phi) + \frac{1}{2}\mu^2 \Phi^\dagger \Phi - \frac{1}{4}\lambda \left(\Phi^\dagger \Phi\right)^2 \\ &\quad + \left(-\lambda_e \bar{E}_L \Phi e_R - \lambda_u \bar{Q}_L \tilde{\Phi} u_R - \lambda_d \bar{Q}_L \Phi d_R + \text{h.c.}\right) \end{aligned} \tag{169}$$

where

$$D_\mu \Phi = \left(\partial_\mu - \frac{1}{2} i g \sigma^i W^i_\mu - \frac{1}{2} g' X_\mu\right) \Phi \tag{170}$$

The total Lagrangian will then be

$$\mathscr{L} = \mathscr{L}_1 + \mathscr{L}_{1/2} + \mathscr{L}_0 \tag{171}$$

There are some important caveats to bear in mind. Our one-generation model, as written above, fails to capture some important features. In particular

- No CP violation can occur in this model. CP violation requires three generations – see e.g. the citation for the 2008 Nobel Prize. Other good resources for grokking this are the particle theory textbook by Cheng and Li and §20 in Peskin and Schroeder.
- Gauge eigenstates for fermions (which transform neatly under isospin and hypercharge gauge transformations) are not necessarily mass eigenstates. In other words, the fermion fields typically get a mass *matrix* from SSB. This complicates the physics significantly, in an interesting way.

3.5 Vector boson masses

The self-interaction we chose for the Higgs included the signature 'wrong-sign' mass term, to trigger spontaneous symmetry breaking.

We pick the vev for the Higgs, spontaneously breaking gauge symmetry, to be aligned as follows:

$$\Phi(x)\Big|_0 = \begin{pmatrix} 0 \\ \eta \end{pmatrix} \tag{172}$$

where

$$\eta \equiv \frac{\mu}{\sqrt{\lambda}} \tag{173}$$

We now impose unitary gauge, in which the charged Higgs is set to zero using local gauge symmetry:

$$\Phi(x) = \begin{pmatrix} 0 \\ \eta + \frac{1}{\sqrt{2}}\sigma(x) \end{pmatrix} \tag{174}$$

To discern which combination of gauge bosons remains massless in response to SSB, we need to look harder at the structure in the $SU(2)$ sector. Note that since $W_\mu = W_\mu^i \left(\frac{1}{2}\sigma^i\right)$ we have, in $SU(2)$ matrix space,

$$W_\mu = \frac{1}{2}\begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} \tag{175}$$

Therefore, as a matrix, the covariant derivative takes the form

$$(D_\mu) = \mathbb{1}\partial_\mu - ig'YX_\mu - \frac{1}{2}ig\begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} \tag{176}$$

This form would be different if we were working with either a different symmetry group or a different matter representation.

Then our Higgs has gauge-covariant derivative

$$D_\mu\Phi = -\frac{1}{2}i\begin{pmatrix} g\eta(W_\mu^1 - iW_\mu^2) + \frac{1}{\sqrt{2}}g\sigma(W_\mu^1 - iW_\mu^2) \\ \sqrt{2}i(\partial\sigma) + \eta(-gW_\mu^3 + g'X_\mu) + \frac{1}{\sqrt{2}}\sigma(-gW_\mu^3 + iX_\mu) \end{pmatrix} \tag{177}$$

It is therefore straightforward to show, as you should work out explicitly, that

$$D^\mu\Phi^\dagger D_\mu\Phi = \frac{1}{2}\partial^\mu\sigma\partial_\mu\sigma + \frac{1}{4}g^2\eta^2\left((W_\mu^1)^2 + (W_\mu^2)^2\right) + \frac{1}{4}\eta^2\left(gW_\mu^3 - g'X_\mu\right)^2 + (\text{cubic}) + (\text{quartic}) \tag{178}$$

We see that *three* vector bosons out of the original four have obtained a mass, while one has not. Specifically, the massive guys are two charged W-bosons

$$W_\mu^\pm \equiv \tfrac{1}{\sqrt{2}}(W_\mu^1 \pm iW_\mu^2) \tag{179}$$

and one neutral boson called the Z:

$$Z_\mu \equiv \frac{(gW_\mu^3 - g'X_\mu)}{\sqrt{g^2 + (g')^2)}} \equiv \cos\theta_W\, W_\mu^3 - \sin\theta_W\, X_\mu \tag{180}$$

By construction, this field is orthogonal to

$$A_\mu \equiv \frac{(g'W_\mu^3 + gX_\mu)}{\sqrt{g^2 + (g')^2}} \equiv \sin\theta_W W_\mu^3 + \cos\theta_W X_\mu \tag{181}$$

The *Weinberg angle* is given by

$$\tan\theta_W = \frac{g'}{g} \tag{182}$$

and so

$$m_W^2 = \frac{1}{2}g^2\eta^2 = \frac{1}{2}g^2\frac{\mu^2}{\lambda} \tag{183}$$

and

$$m_Z^2 = \frac{m_W^2}{\cos^2\theta_W} > m_W^2 \tag{184}$$

Working through the algebra gives, for the fermion-gauge interaction Lagrangian,

$$\begin{aligned}
\mathscr{L} \supset\; & g\sin\theta_W \left[(-1)\bar{e}\gamma^\mu e + (\tfrac{2}{3})\bar{u}\gamma^\mu u + (-\tfrac{1}{3})\bar{d}\gamma^\mu d\right] A_\mu \\
& + \frac{g}{\cos\theta_W}\left[+\bar{\nu}_e\gamma^\mu(\tfrac{1}{2})\nu_e + \bar{e}_R\gamma^\mu(\sin^2\theta_W)e_R + \bar{u}_R\gamma^\mu(-\tfrac{2}{3}\sin^2\theta_W)u_R + \bar{d}_R\gamma^\mu(+\tfrac{1}{3}\sin^2\theta_W)d_R\right. \\
& \left.+\bar{e}_L\gamma^\mu(-\tfrac{1}{2}+\sin^2\theta_W)e_L + \bar{u}_L\gamma^\mu(+\tfrac{1}{2}-\tfrac{2}{3}\sin^2\theta_W)u_L + \bar{d}_L\gamma^\mu(-\tfrac{1}{2}+\tfrac{1}{3}\sin^2\theta_W)d_L\right] Z_\mu \\
& + \frac{g}{\sqrt{2}}\left[(\bar{\nu}_e\gamma^\mu e_L + \bar{u}_L\gamma^\mu d_L)\, W_\mu^+ + \text{h.c.}\right]
\end{aligned} \tag{185}$$

The vector boson A_μ couples only to electrons, not to neutrinos. This is of course exactly what we wanted: electrons are charged while neutrinos are electrically neutral. Even the form of the "electromagnetic current"

$$J_\mu^{\text{EM}} = (-1)\bar{e}\gamma_\mu e + (\tfrac{2}{3})\bar{u}\gamma_\mu u - (\tfrac{1}{3})\bar{d}\gamma_\mu d \tag{186}$$

is correct, if we identify

$$e = g\sin\theta_W \tag{187}$$

The neutral Z_μ boson couples to the "neutral current", with strength $g/\cos\theta_W$

$$\begin{aligned}
J_\mu^{\text{n}} &= \bar{\nu}_e\gamma_\mu(\tfrac{1}{2})\nu_e + \bar{e}_R\gamma_\mu(\sin^2\theta_W)e_R + \bar{u}_R\gamma_\mu(-\tfrac{2}{3}\sin^2\theta_W)u_R + \bar{d}_R\gamma_\mu(+\tfrac{1}{3}\sin^2\theta_W)d_R \\
&+ \bar{e}_L\gamma_\mu(-\tfrac{1}{2}+\sin^2\theta_W)e_L + \bar{u}_L\gamma_\mu(+\tfrac{1}{2}-\tfrac{2}{3}\sin^2\theta_W)u_L + \bar{d}_L\gamma_\mu(-\tfrac{1}{2}+\tfrac{1}{3}\sin^2\theta_W)d_L
\end{aligned} \tag{188}$$

This talks to neutrinos, electrons and quarks. It is not left-right symmetric.

The "charged current" for the W_μ^+, coupling to it with strength $g/\sqrt{2}$ is

$$J_\mu^{\text{c}} = (\bar{\nu}_e\gamma_\mu e_L + \bar{u}_L\gamma_\mu d_L) \tag{189}$$

and similarly for its Hermitean conjugate for the W_μ^-. This vertex is relevant for e.g. beta decay.

3.6 Fermion masses

Recall that we started out with a Lagrangian with no explicit fermion masses, in order to be consistent with $SU(2) \times U(1)$ gauge invariance. We could however write down a gauge-invariant coupling between the Higgs and the fermions that took the form of a Yukawa coupling

$$\begin{aligned}
\mathscr{L}_{\text{Yuk}} &\supset -f\left(\bar{E}_L\Phi e_R + \bar{e}_R\Phi^\dagger E_L\right) + (\text{quarks}) \\
&= -\lambda_e\left(\bar{\nu}_e e_R\phi^+ + \bar{e}_L e_R\phi^0 + \bar{e}_R\nu_e(\phi^+)^\dagger + \bar{e}_R e_L(\phi^0)^\dagger\right) + (\text{quarks})
\end{aligned} \tag{190}$$

In unitary gauge, when SSB occurs, this single-generation Lagrangian collapses to

$$\mathscr{L}_{\text{Yuk}} \supset -\lambda_e\eta\,(\bar{e}_L e_R + \bar{e}_R e_L) + (\text{cubic}) + (\text{quarks}) \tag{191}$$

One very important consequence of this Standard Model – just as true for three generations as for one – is that neutrino masses are identically zero.

$$m_\nu^2 \equiv 0 \tag{192}$$

There is simply no way to create neutrino masses consistent with $SU(2) \times U(1)$ gauge invariance. Accordingly, all neutrinos are of left-handed chirality while all antineutrinos are right-handed.

The electrons and quarks, on the other hand, have masses governed by the strength of the Yukawa couplings $\lambda_{e,u,d}$ as well as the vev of the Higgs, for instance

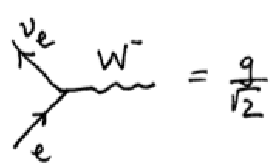
$$m_e = (\lambda_e)\frac{\mu}{\sqrt{\lambda}} \tag{193}$$

This expression for the fermion mass is all well and good, but how does it relate to quantities that we can actually measure in the laboratory? In order to tease that apart, it helps to recall

$$m_W = \frac{1}{\sqrt{2}} g \frac{\mu}{\sqrt{\lambda}} \tag{194}$$

Therefore, to find the ratio of the electron mass to the W mass, we need to focus on relating the strength of the Yukawa coupling λ_e to the $SU(2)$ gauge coupling g.

The charged current coupling to $W^\pm$ is simpler in structure than the neutral current, so let us focus there. We have $g/\sqrt{2}$ for either the $\bar{\nu}_e$–e_L–W^+ vertex or for the $\bar{u}_L$–d_L–W^+ vertex, and the same goes for their Hermitean conjugates.



Consider the structure of the propagator for the massive gauge bosons. At low momentum, the detailed momentum dependence in that propagator, whatever it is in whichever gauge, can be expected to collapse into the simple expression

$$\lim_{k^2 \to 0} \Delta^{(W)}_{\alpha\beta}(k) = \frac{i\eta_{\alpha\beta}}{m_W^2} \tag{195}$$

Combined with this propagator, the Yukawa couplings would give rise to an effective four-Fermi interaction at low energy of the form

$$\mathscr{L}_{\mathrm{W}} \supset -\frac{g^2}{2} J^c_\mu(\text{quark})^\dagger \frac{\eta^{\mu\nu}}{m_W^2} J^c_\nu(\text{lepton}) \tag{196}$$

where

$$\begin{aligned} J^c_\mu(\text{lepton}) &= \bar{\nu}_e \gamma_\mu e_L \\ J^c_\mu(\text{quark}) &= \bar{d}_L \gamma_\mu u_L \end{aligned} \tag{197}$$

The coefficient of the effective four-Fermi interaction was historically written as the *Fermi decay constant* of Enrico Fermi's attempt to explain the weak interactions. At tree level,

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} \tag{198}$$

Very similar logic can be used to obtain the four-fermi vertex for muon decay (into an electron and some neutrinos) as well.

The experimentally measured value is

$$G_F \simeq \frac{10^{-5}}{m_{\text{proton}}^2} \tag{199}$$

From this, we can conclude (modulo renormalization!) that the $SU(2)$ coupling is

$$g^2 \simeq 4\sqrt{2} m_W^2 G_F \simeq \frac{10^{-5} m_W^2}{m_{\text{proton}}^2} \tag{200}$$

This is a number less than unity, but is *not especially weak*. It is really the size of the W boson mass that makes weak interactions at low energy look so weak, not so much the weakness of g^2 itself.

Neutral current scattering mediated by the Z_μ can give rise to neutrino-electron (or mu, or tau) scattering which is of similar order as the charged current scattering process as long as the Weinberg angle is not too small. Working along similar lines to our method above, we can actually find the effective four-fermi interaction Lagrangian for both W and Z exchange at low energy. As you should check for your own satisfaction, this yields

$$\lim_{k^2 \to 0} (\Delta \mathscr{L}_W + \Delta \mathscr{L}_Z) = \frac{4G_F}{\sqrt{2}} \left[(J_\mu^1)^2 + (J_\mu^2)^2 + (J_\mu^3 - \sin^2\theta_W J_\mu^{\text{EM}})^2 \right] \tag{201}$$

This expression, as Peskin and Schoeder explicitly point out, becomes manifestly invariant under an unbroken *global* $SU(2)$ symmetry in the limit that $g' \to 0$ or $\sin^2\theta_W \to 0$. This custodial symmetry demands that $m_Z = m_W$ and has some interesting physical consequences that we do not have time to discuss here. For more details, see Chapter 20 of Peskin and Schroeder. Note: any theory whose Lagrangian possesses a global $SU(2)$ symmetry gives rise to neutral current processes; you only get the intermediate vector bosons arising if the symmetry is gauged.

The Fermi theory of beta decay was a decent low-energy effective theory. However, it has one inexcusable drawback for anyone wanting to discover the deep structure of quark and lepton interactions: it is non-renormalizable. Like Einstein's theory of GR, the quantum Fermi theory has ultraviolet infinities that cannot be mathematically tamed and are not physically reasonable. This is a big reason why we all teach the Standard Model in terms of spontaneously broken $SU(2) \times U(1)$ gauge symmetry instead – this gauge theory *is* renormalizable.

Is the Glashow-Weinberg-Salam model (for which they shared the Nobel Prize) a consistent quantum theory? One topic which we may have time to develop in the final week of this course is the topic of *anomalies*. Anomalies arise from loop diagrams generically, and may spoil gauge invariance by breaking it at quantum loop level. In a beautiful twist of group theory, the charge assignments we have made and the choice of gauge group give rise to cancellation of all anomalies - provided that our chiral fermions appear in complete doublet generations. This is one way that the existence of both the charm and top quark were predicted: physics was known to be inconsistent for the existing odd number of quarks.

So how should we think about this lovely construction we have just spent an entire week discussing, in the context of "Unification"? In the history of the universe since the Big Bang (or Big Brane Crash, or whatever it was!), the ambient temperature has plummeted as our universe underwent great expansion. CMBR photons today sit in the microwave, but were a great deal hotter during the very early universe. Thermal fluctuations of a typical quantum field in equilibrium are driven by the ambient temperature. At high-T, the $|\Phi|^4$ part of the Higgs potential will dominate, and the little bump of a local minimum at $\Phi = 0$ will hardly be noticed. All four vector bosons of $SU(2) \times U(1)$ are massless in this phase. At lower temperatures, however, the average thermal energy for any field mode gets drastically reduced, and so the topography of the Mexican Hat potential will become very important. Once the average thermal energy drops below the height of the bump, SSB occurs. We see that at low energy $W^{\pm}, Z$ become massive – as do all the quarks and leptons save neutrinos – while the photon stays exactly massless.

3.7 Multiple generations and CP violation

Before we move on to Feynman path integrals, we can give a quick outline of why CP violation in the Standard Model requires three generations.

Consider 3 generations of quarks and leptons as fundamental vectors in flavour space:

$$(e'_i) = \begin{pmatrix} e' \\ \mu' \\ \tau' \end{pmatrix} \qquad (\nu'_i) = \begin{pmatrix} \nu'_e \\ \nu'_\mu \\ \nu'_\tau \end{pmatrix} \qquad (p'_i) = \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} \qquad (n'_i) = \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} ,$$

where $i = 1 \ldots 3$. These can be combined into $SU(2)$ isospin doublets in gauge space as

$$L'_{iL} = \begin{pmatrix} \nu'_i \\ e'_i \end{pmatrix}_L \qquad Q'_{iL} = \begin{pmatrix} p'_i \\ n'_i \end{pmatrix}_L .$$

Such *gauge eigenstates* transform nicely under weak isospin and hypercharge.

Starting from minimal coupling and known isospins and hypercharges, we can show that the interaction Lagrangian between the above fermions and the vector bosons $\{W^A, X\}$ is

$$\begin{aligned} \mathscr{L}_{\rm min} = \bar{L}'_{iL} i \left(\not{\partial} - i\frac{g}{2}\sigma^A \not{W}^A + i\frac{g'}{2}\not{X} \right) L'_{iL} + \bar{e}'_{iR} i \left(\not{\partial} + ig'\not{X} \right) e'_{iR} + \\ + \bar{Q}'_{iL} i \left(\not{\partial} - i\frac{g}{2}\sigma^A \not{W}^A - i\frac{g'}{6}\not{X} \right) Q'_{iL} + \bar{p}'_{iR} i \left(\not{\partial} - \frac{2i}{3}g'\not{X} \right) p'_{iR} + \bar{n}'_{iR} i \left(\not{\partial} + \frac{i}{3}g'\not{X} \right) n'_{iR} . \end{aligned}$$

Using symmetry arguments exactly patterned on the one-generation case we covered earlier, we can also see why Yukawa couplings for the multi-generation case must take the form

$$\mathscr{L}_{\rm Yuk} = f^{[e]}_{ij}\, \bar{L}'_{iL}\, \Phi\, e'_{jR} + f^{[u]}_{ij}\, \bar{Q}'_{iL}\, \tilde{\Phi}\, p'_{jR} + f^{[d]}_{ij}\, \bar{Q}'_{iL}\, \Phi\, n'_{jR} ,$$

where $f_{ij}^{[e,u,d]}$ are Yukawa couplings for leptons, up-type quarks and down-type quarks, and

$$\tilde{\Phi} = \frac{i}{2}\sigma_2 \Phi^* \,.$$

Recall that this $\tilde{\Phi}$ is built entirely out of Φ; it is not a new Lagrangian field.

For the rest of this discussion we will focus on the quark sector only, for simplicity.

Now imagine SSB happening. There will generically be a mass *matrix* M for quarks, as you can see by looking at the form of $\mathscr{L}_{\rm Yuk}$. There is no physical reason to expect the quark mass matrix to be diagonal - or even symmetric or Hermitean. This mass matrix M can be diagonalized by a biunitary transformation

$$S^\dagger M T = M_d \,,$$

where M_d is diagonal and S, T are unitary. Accordingly, we can show that for a quark gauge eigenstate ψ' there is a ψ which is a quark mass eigenstate:

$$\bar{\psi}'_L M \psi'_R = \bar{\psi}_L M_d \psi_R \,,$$

where

$$\psi'_L = S\psi_L \text{ and } \psi'_R = T\psi_R \,.$$

Suppose that we have n generations of quarks, each of which is an isospin doublet. A general $n \times n$ unitary matrix U has n^2 real components, and this can be characterized in terms of $n(n-1)/2$ real rotation angles and $n(n+1)/2$ complex phases. However, not all these phases turn out to be physically observable. Some of them can be eliminated by quark field redefinitions! In particular, if we note the fact that only *charged* weak currents are nondiagonal in quark flavours,

$$\mathscr{L}_{\rm CC} = \frac{g}{\sqrt{2}} \left(\bar{u}, \bar{c}, \bar{t}\right)_L \gamma^\mu \; U \; \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L W_\mu^+ + \text{h.c.}$$

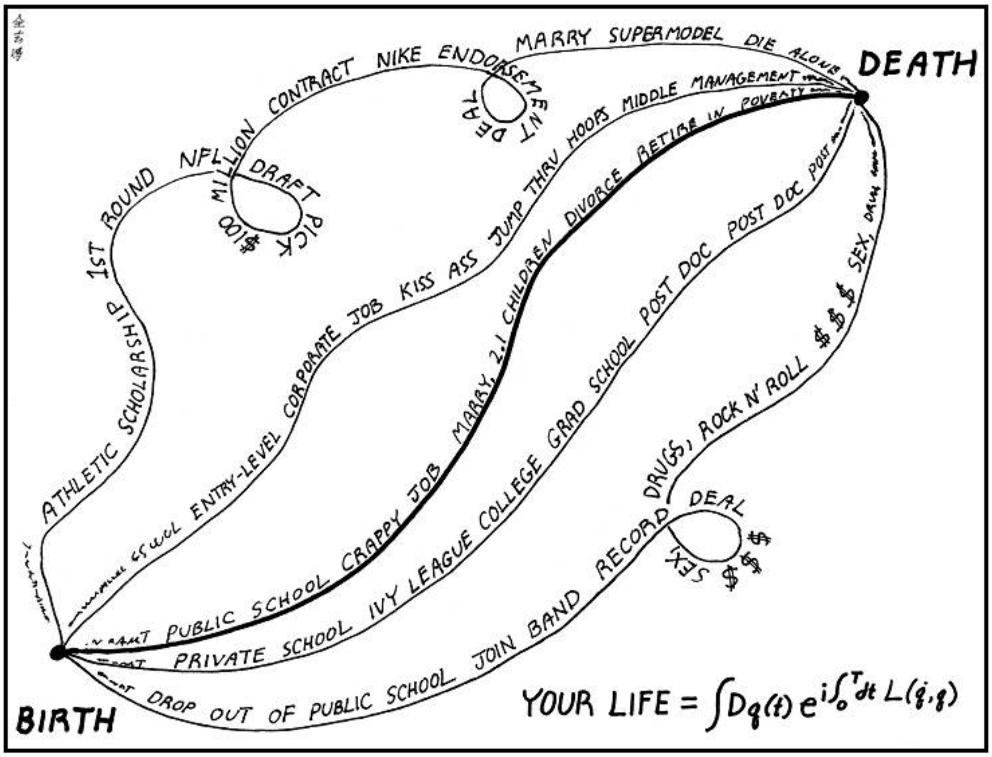
we can see that changing the phase of only one row or column of a unitary matrix appears to not change any physical observables. There are $2n$ such phases. But we actually double-counted one: pulling out a general phase out front of the whole unitary matrix does not change physical observables either. So there are actually only $2n-1$ unphysical phases. Ergo, the final number of physically observable phases is

$$\frac{1}{2}n(n+1) - (2n-1) = \frac{1}{2}(n-1)(n-2) \,. \tag{202}$$

We therefore need at least 3 generations for CP violation: $n = 1$ or $n = 2$ won't cut it.

4 The Feynman Path Integral

4.1 FPI for non-relativistic point particles



The Path Integral Formulation of Your Life

Note: the discussion in this subsection is based on §9 of Peskin and Schroeder.

Recall from QFT1 that if we wanted to compute the quantum mechanical amplitude for a particle to go from x_a to x_b in a time T, we would write down the expression

$$U(x_a, x_b; T) = \langle x_b | e^{-iHT} | x_a \rangle \, . \tag{203}$$

where H is the time evolution operator. Richard Feynman, a truly great theorist *and* experimentalist, proposed that we should think about quantum mechanics as a sum over all possible paths. Specifically, Feynman proposed that

$$U(x_a, x_b; T) = \sum_{\text{all paths}} \exp\left(i[\text{phase}]\right) \, , \tag{204}$$

where the phase is proportional to the action S. This is motivated because classical paths have stationary phase and stationary action,

$$\frac{\delta}{\delta x} S \bigg|_{x = x_{\text{cl}}} = 0 \, . \tag{205}$$

The role of the proportionality constant between phase and action is played, as you might expect from dimensional analysis, by $\hbar$. In other words, Feynman proposed that

$$\langle x_b, x_a; T\rangle = \int \mathscr{D}x(t) \exp\left(\frac{iS[x(t)]}{\hbar}\right) , \tag{206}$$

for *very* general systems.

To see how the above **functional integral** works, in particular to understand $\mathscr{D}x(t)$, it is instructive to figure out the measure by discretizing. This discretization trick is extremely useful in quantum field theory in general, and in fact all of lattice gauge theory is based on it. Although, when a physical theory is properly understood, the same result must of course be obtained by any other regularization method.

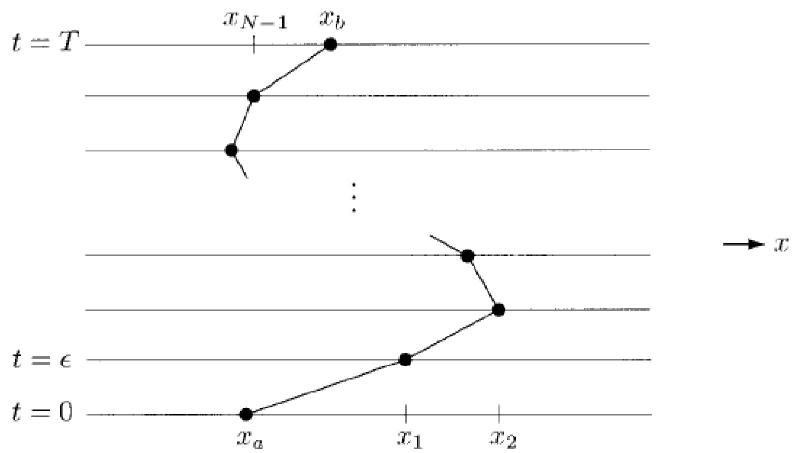


Figure 9.2. We define the path integral by dividing the time interval into small slices of duration ϵ, then integrating over the coordinate x_k of each slice.

Consider time evolution from $t = 0$ to $t = T$ in steps of ϵ, approximating the paths by a sequence of straight lines. Since we are working in the non-relativistic approximation for a point particle, we have for the action

$$S = \int_0^T dt \left(\frac{1}{2}m\dot{x}^2 - V(x)\right) \rightarrow \sum_k \left[\frac{m}{2\epsilon}(x_{k+1} - x_k)^2 - \epsilon V\left(\frac{x_{k+1} + x_k}{2}\right)\right] . \tag{207}$$

Anticipating a future need, we allow each segment to have its own normalization factor in the definition of the measure, and write

$$\begin{aligned}\int \mathscr{D}x(t) &= \frac{1}{C(\epsilon)} \int \frac{dx_1}{C(\epsilon)} \int \frac{dx_2}{C(\epsilon)} \cdots \int \frac{dx_{N-1}}{C(\epsilon)} \\ &= \frac{1}{C(\epsilon)} \prod_k \int_{-\infty}^{+\infty} \frac{dx_k}{C(\epsilon)} .\end{aligned} \tag{208}$$

Note that one factor of $C(\epsilon)$ was included for each of the N time slices. Our task here is to find, in the continuum limit $\epsilon \rightarrow 0$, the factors $C(\epsilon)$ and the differential equation obeyed by our friend U.

Consider the difference between the penultimate and final steps in the sequence $x_1, x_2, \cdots x_N$. For this step we have

$$U(x_a, x_b; T) = \int_{-\infty}^{+\infty} \frac{dy}{C(\epsilon)} \exp\left(\frac{i}{\hbar}\left[\frac{m(x_b-y)^2}{2\epsilon} - \frac{i}{\hbar}\epsilon V(\frac{x_b+y}{2})\right]\right) \times U(x_a, y; T-\epsilon)\,, \quad (209)$$

where the right factor of $U(x_a, y; T-\epsilon)$ incorporates all the data from the previous slices.

Now notice how this integrand behaves in the limit $\epsilon \to 0$. First, note that the potential term varies weakly with ϵ, as $V(x)$ is well behaved. Secondly, the kinetic term oscillates wildly in the phase $e^{iS/\hbar}$, which makes the energy cost prohibitive unless y is kept really close to x_b. Therefore, expanding in a Taylor series, we have

$$\begin{aligned} U(x_a, x_b; T) &= \int_{-\infty}^{+\infty} \frac{dy}{C(\epsilon)} \exp\left(\frac{i}{\hbar}\frac{m}{2\epsilon}(x_b-y)^2\right) \times \left[1 - \frac{i\epsilon}{\hbar}V(x_b) + \cdots\right] \times \\ &\quad \times \left[1 + (y-x_b)\frac{\partial}{\partial x_b} + \frac{1}{2}(y-x_b)^2\frac{\partial^2}{\partial x_b^2} + \cdots\right] \times U(x_a, x_b; T-\epsilon)\,. \quad (210) \end{aligned}$$

Because of the Taylor expansion, the integral over y reduces to a simple Gaussian. Recalling that

$$\int_{-\infty}^{+\infty} dy\, e^{-by^2} = \sqrt{\frac{\pi}{b}}\,, \qquad \int_{-\infty}^{+\infty} dy\, y^2 e^{-by^2} = \frac{1}{2b}\sqrt{\frac{\pi}{b}}\,, \quad (211)$$

we have that

$$U(x_a, x_b; T) = \left(\frac{1}{C}\sqrt{\frac{\epsilon\, 2\pi\hbar}{-im}}\right)\left\{1 - \frac{i\epsilon}{\hbar}V(x_b) + \frac{i\epsilon\hbar}{2m}\frac{\partial^2}{\partial x_b^2} + \mathscr{O}(\epsilon)^2\right\} \times U(x_a, x_b; T-\epsilon)\,. \quad (212)$$

This equation does not make any sense as $\epsilon \to 0$, *unless*

$$C(\epsilon) = \sqrt{\frac{\epsilon\, 2\pi\hbar}{-im}}\,. \quad (213)$$

This requirement then implies the following differential equation for U in the continuum limit:

$$i\hbar\frac{\partial}{\partial T}U(x_a, x_b; T) = \left\{-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x_b^2} + V(x_b)\right\} U(x_a, x_b; T)\,, \quad (214)$$

which is our old friend the Schrödinger equation!

One more consistency check is in order. We know from our adventures with canonical quantization that

$$\lim_{T\to 0}\langle x_b|e^{-iHT}|x_a\rangle = \delta(x_a - x_b)\,. \quad (215)$$

Compare this to our discretized equation,

$$\sqrt{\frac{-im}{2\pi\hbar\epsilon}}\exp\left(\frac{i}{\hbar}\frac{m}{2}(x_b-x_a)^2 + \mathscr{O}(\epsilon)\right)\,. \quad (216)$$

The discretized quantity, in the continuum limit $\epsilon \to 0$, is exactly the same mathematical beast as the delta function. We just regularized it in such a way that it is the limit of a Gaussian in the discretized picture.

This is excellent – our Feynman Path Integral story and canonical quantization give the same result. Fortunately, this agreement between functional and canonical quantization methods is not limited to a simple non-relativistic point particle. In all QFTs understood well enough to analyze, this equivalence remains solid. The reason we are switching you to the path integral is that beyond spin-half it is *much* easier to calculate with – especially if you have any meaningful symmetry in the picture such as gauge symmetry.

How about the extension to any system with (q^i, p_j)? The analysis is very similar, except in a few places. First, a useful identity is $1 = \left(\prod_i dq_i^k\right) |q^k\rangle\langle q^k|$, where k indexes timesteps and i indexes coordinates q^i. Second, as Peskin and Schroeder explain on p.281 of their textbook, for general *operator* ps and qs it is ambiguous as to how to define matrix elements we need. For brevity's sake, it is assumed that the Hamiltonian is *Weyl ordered*, to remove position/momentum ordering ambiguities. The next step in the story involves expanding in momentum eigenstates and position/momentum wavefunction overlaps. For the cases in which the Hamiltonian is a function of only the canonical momenta, the algebra goes through straightforwardly. More details may be found in Peskin and Schroeder. After the algebraic dust settles, the result in the continuum limit is

$$U(q_0, q_N; T) = \left\{ \prod_i \int \frac{dq^i dp_i}{2\pi\hbar} \right\} \exp\left(\frac{i}{\hbar} S\right) . \tag{217}$$

Of course, the measure appearing in this continuum expression is none other than the usual measure on phase space. This is what we will now generalize to fields.

We can now take this formula derived for dynamical coordinates $q^a(t)$ and canonical momenta $p_b(t)$ and realize that the only difference for quantum fields – animals that are functions of relativistic coordinates x^μ – will be to replace

$$\begin{aligned} t &\longrightarrow x^\mu \\ q^i(t) &\longrightarrow \Phi^A(x^\mu) \, , \end{aligned} \tag{218}$$

where A is a collection of spacetime and/or internal indices on the quantum field.

4.2 Functional quantization for scalar fields

The functional quantization method, as compared to the canonical approach of expanding field operators in Fourier expansions of creation and annihilation operators, has a number of distinct advantages. One of the most important is that the Feynman Path Integral (FPI) inherently preserves symmetries manifest in the action such as gauge symmetry and Lorentz symmetry – something the canonical approach cannot do.

Let us now focus on a specific example of how to do field quantization in the functional approach. We pick the spin zero scalar field to begin with, as it lets us see the physics without getting distracted by spinors or gauge invariance. We choose a minimal kinetic term and write

$$\mathscr{L} = \tfrac{1}{2}\partial^\mu\phi\partial_\mu\phi - \tfrac{1}{2}m^2\phi^2 - V(\phi) \, . \tag{219}$$

Recall that this gave rise to a Feynman propagator

$$\text{———} \qquad \frac{i}{(k^2 - m^2 + i\epsilon)}$$

To see how the Lagrangian arises in the phase for scalar field theory, we need to work out the transition amplitude. To find $e^{-iHT/\hbar}$, we need the Hamiltonian density, which is

$$\begin{aligned}\mathscr{H} &= \Pi \cdot \dot{\Phi} - \mathscr{L} \\ &= \tfrac{1}{2}\Pi^2 + \tfrac{1}{2}|\nabla\phi|^2 + \tfrac{1}{2}m^2\phi^2 + V(\phi)\,,\end{aligned} \tag{220}$$

where Π is the field canonical momentum $\partial\mathscr{L}/\partial\dot{\phi}$. Then

$$\begin{aligned}\langle\phi_b(x)|e^{-iHT/\hbar}|\phi_a(x)\rangle &= \int \mathscr{D}\phi\mathscr{D}\Pi \exp\biggl(\frac{i}{\hbar}\int_0^T dt \int d^dx \,\times \\ &\times \Bigl[\Pi\cdot\dot{\phi} - \tfrac{1}{2}\Pi^2 - \tfrac{1}{2}|\nabla\phi|^2 - \tfrac{1}{2}m^2\phi^2 - V(\phi)\Bigr]\biggr)\,.\end{aligned} \tag{221}$$

Since this is an integral at most quadratic in Π, we can complete the square and do the integral analytically: it is just a Gaussian. (If there's anything a $\hbar$ physicist can do, it's Gaussian integrals!) The result just gives a normalization constant out front, allowing us to define

$$\langle\phi_b(x)|e^{-iHT/\hbar}|\phi_a(x)\rangle = \int \mathscr{D}\phi \exp\left(\frac{i}{\hbar}\int_0^T dt \int d^dx\, \mathscr{L}[\phi(x^\mu)]\right)\,. \tag{222}$$

N.B.: Morally, you should think of the Hamiltonian as being defined via this relationship. It is the Feynman path integral – the right hand side of this equation – that we will consider fundamental to Quantum Field Theory. FPI technology is extremely powerful and will allow us to quantize fields of spin one and greater with a lot more ease than with canonical apparatus.

Physicists studying QFTs always want to know about how field operators correlate amongst one another. The most basic correlator to study is the **two-point correlation function**, and so we now set up the procedure for finding it.

Suppose we consider time to run from $-T$ to $+T$, and that t_1, t_2 are two intermediate times in the range $[-T, T]$. Suppose further that we define our scalar field to take endpoint values

$$\begin{aligned}\phi(-T, x) &= \phi_a(x) \\ \phi(+T, x) &= \phi_b(x)\,,\end{aligned} \tag{223}$$

and let the interior values at t_1, t_2 be defined as

$$\begin{aligned}\phi(t_1, x) &= \phi_1(x) \\ \phi(t_2, x) &= \phi_2(x)\,.\end{aligned} \tag{224}$$

Consider the following quantity

$$\Xi := \int \mathscr{D}\phi(x)\,\{\phi(x_1)\phi(x_2)\}\exp\left(i\int_{-T}^{+T}\mathscr{L}(\phi)\right)\,. \tag{225}$$

We will use this as a starting point and then take a specific type of large-T limit in order to extract the two-point correlation function.

First, let us split the measure as follows:

$$\int \mathscr{D}\phi(x) = \int \mathscr{D}\phi_1(x)\mathscr{D}\phi_2(x) \int_{\substack{\phi(t_1,x)=\phi_1(x) \\ \phi(t_2,x)=\phi_2(x)}} \mathscr{D}\phi(x)\,. \tag{226}$$

Now let us suppose that $t_1 < t_2$. Then

$$\begin{aligned} \Xi_{t_1<t_2} &= \int \mathscr{D}\phi_1(x)\mathscr{D}\phi_2(x)\,\{\phi_1(x_1)\phi_2(x_2)\}\times \\ &\quad \times\langle\phi_b|e^{-iH(T-t_2)/\hbar}|\phi_2\rangle\langle\phi_2|e^{-iH(t_2-t_1)/\hbar}|\phi_1\rangle\langle\phi_1|e^{-iH(t_1+T)/\hbar}|\phi_a\rangle\,. \end{aligned} \tag{227}$$

To simplify this we need only recall that, for a quantum field ϕ, $\phi(x)$ is the eigenvalue obtained when the Schrödinger picture field operator hits a field state:

$$\hat{\phi}_S|\phi_1\rangle = \phi_1(x_1)|\phi_1\rangle\,. \tag{228}$$

Also, recall the completeness relation for the quantum field ϕ:

$$\int \mathscr{D}\phi\,|\phi\rangle\langle\phi| = \mathbb{1}\,. \tag{229}$$

Using that fact, we get

$$\Xi_{t_1<t_2} = \langle\phi_b|e^{-iH(T-t_2)/\hbar}\hat{\phi}_S(x_2)e^{-iH(t_2-t_1)/\hbar}\hat{\phi}_S(x_1)e^{-iH(t_1+T)/\hbar}|\phi_a\rangle\,. \tag{230}$$

Notice something interesting: the quantity $e^{iHt}\hat{\phi}_S(x)e^{-iHt/\hbar}$ is just the Heisenberg picture field operator $\hat{\phi}_H(x)$. Therefore,

$$\Xi_{t_1<t_2} = \langle\phi_b|e^{-iHT/\hbar}\phi_H(t_2,x_2)\phi_H(t_1,x_1)e^{-iHT/\hbar}|\phi_a\rangle\,. \tag{231}$$

The expression for $t_2 < t_1$ is identical, except for a shuffling of indices.

Our final trick is to strip off the T-dependence in order to get an unsullied two-point correlation function. Assuming that our states have any overlap with the vacuum $|\Omega\rangle$, we take a specific kind of long-time limit:

$$T \to \infty(1-i\epsilon)\,. \tag{232}$$

The reason for introducing the $i\epsilon$ is mathematical: to ensure convergence. Physically, what this long-time limit does is project the vacuum $|\Omega\rangle$ from $|\phi_a\rangle$ and project $\langle\Omega|$ from $\langle\phi_a|$. To see how this works, we simply decompose $|\phi_a\rangle$ into an energy eigenbasis $|n\rangle$ of $\hat{H}$, and consider how it behaves in the long-time limit specified above. We find

$$\begin{aligned} e^{-iHT/\hbar}|\phi_a\rangle &= \sum_n e^{-iE_nt/\hbar}|n\rangle\langle n|\phi_a\rangle \\ &\xrightarrow[T\to\infty(1-i\epsilon)]{} \langle\Omega|\phi_a\rangle e^{-iE_0\cdot\infty\cdot(1-i\epsilon)/\hbar}|\Omega\rangle\,. \end{aligned}$$

It is easy to see from this expression why we chose the limit $T \to \infty(1-i\epsilon)$: only for this sign choice does the ϵ term give a damping and hence a physically reasonable regularization. It corresponds physically to using the Feynman propagator (half-advanced half-retarded) with the under-and-over contour as depicted below.

To make sense of the infinitely oscillating phase part, we simply compute a *ratio* as follows:

$$\langle\Omega|T\left\{\phi_H(x_1)\phi_H(x_2)\right\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathscr{D}\phi \cdot \phi(x_1)\phi(x_2)\cdot \exp\left(i\int_{-T}^{+T} d^Dx\,\mathscr{L}\right)}{\int \mathscr{D}\phi \exp\left(i\int_{-T}^{+T} d^Dx\,\mathscr{L}\right)}\,. \tag{233}$$

This final expression is a very important formula, to which we will return repeatedly. As you can see for yourself, it easily generalizes to higher correlation functions.

4.3 Doing the FPI for a free massive scalar field via spacetime discretization

In a general case, the FPI cannot be done exactly. Most of the time, and for all of this course, it has to be approached perturbatively. For clarity we now show explicitly how to compute the path integral for a *free* massive scalar field (with quadratic kinetic term) in detail, by introducing a discretization of spacetime.

We pick spacetime to be a square lattice of spacing ϵ . We further take the volume of d-dimensional space to be $V = L^d$ (effectively an infrared cutoff). Later on we will take the $\epsilon \to 0$ and $V \to \infty$ limits in order to recover the continuum limit. Expanding our quantum field possessing Klein-Gordon action in terms of Fourier modes gives

$$\phi(x_i) = \frac{1}{V}\sum_n e^{-ik_n\cdot x_i}\phi(k_n) \tag{234}$$

where the wavevector k_n^μ is

$$k_n^\mu = \frac{2\pi n^\mu}{L} \tag{235}$$

where

$$n^\mu \in \mathbb{Z}\,, \qquad |k^\mu| < \pi/\epsilon \tag{236}$$

(Note that, in the continuum limit, $\frac{1}{V}\sum_n \longrightarrow \int \frac{d^Dk}{(2\pi)^D}$.)

In general the Fourier coefficients are complex; for a real scalar field we have

$$\phi^\star(k) = \phi(-k) \tag{237}$$

Instead of working with ϕ and $\phi^\star$ we can just as easily change basis to work with the real and imaginary parts of $\phi(k_n)$, namely $\Re\phi(k_n), \Im\phi(k_n)$, where $k_n^0 > 0$. Since a basis change is just a unitary transformation, we can write the measure as

$$\mathscr{D}\phi(x) = \prod_{k_n^0>0} d\Re\phi(k_n)d\Im\phi(k_n) \tag{238}$$

In this approach where we discretized spacetime on a square lattice, we have for the scalar field action ($D = d+1$)

$$\begin{aligned} S_0 &= \int d^D x \left(\tfrac{1}{2}\partial^\mu \phi \partial_\mu \phi - \tfrac{1}{2} m^2 \phi^2\right) \\ &= \frac{1}{V}\sum_{k_n^0 > 0} \tfrac{1}{2}\left(k_n^2 - m^2\right)|\phi(k_n)|^2 \\ &= -\frac{1}{V}\sum_{k_n^0>0}(m^2 - k_n^2)\left[\Re\phi_n{}^2 + \Im\phi_n{}^2\right] \end{aligned}$$

Notice that for our real scalar field $m^2 - k_n^2 = m^2 - |\vec{k_n}|^2 + (k_n^0)^2$, a quantity which is positive as long as k_n^0 is large enough. We will assume that this is satisfied in the following; if not, we will define it in other regions by analytic continuation.

Now we can actually evaluate our Feynman path integral for our free real scalar field in the discretized spacetime regularization:

$$\begin{aligned} Z_{s=0,\text{free}} &= \int \mathscr{D}\phi e^{iS_0} \\ &= \left(\prod_{k_n^0>0}\int d(\Re\phi_n) d(\Im\phi_n)\right) \exp\left(-\frac{i}{V}\sum_{n|k_n^0>0}(m^2 - k_n^2)|\phi_n|^2\right) \\ &= \prod_{k_n^0>0}\int d(\Re\phi_n) d(\Im\phi_n) \exp\left(-\frac{i}{V}\sum_{n|k_n^0>0}(m^2 - k_n^2)|\Re\phi_n|^2\right) \times \\ &\quad \times \exp\left(-\frac{i}{V}\sum_{n|k_n^0>0}(m^2 - k_n^2)|\Im\phi_n|^2\right) \\ &= \prod_{k_n^0>0}\sqrt{\frac{-i\pi V}{(m^2 - k_n^2)}}\sqrt{\frac{-i\pi V}{(m^2 - k_n^2)}} \\ &= \prod_{k_n}\sqrt{\frac{-i\pi V}{(m^2 - k_n^2)}} \end{aligned}$$

Given the prevalence of $(m^2 - k_n^2)$ in the above expressions, we should wonder how the contour in the complex plane should be handled. Earlier, we needed to take the limit $T \to \infty(1 - i\epsilon)$ in order to make a well-defined path integral. Now, note that rotating a contour $90^\odot$ counterclockwise in the complex plane yields $t \to t(1 - i\epsilon)$. Accordingly, k^0 – with its upstairs index – transforms as $k^0 \to k^0(1 - i\epsilon)$ in all expressions. Therefore,

$$(k^2 - m^2) \to (k^2 - m^2 + i\epsilon) \tag{239}$$

This contour choice is *necessary* in order to ensure that the Gaussian integrals making up Z converge to a physically sensible answer.

As yet, our final expression for Z is not written such a way that the generalization to non-free fields is obvious. We now make a quick detour to describe another way of writing the answer, in terms of what is known as a **functional determinant**, which makes the generalization to interactions and other spins more clear.

Consider a Gaussian integral of the form

$$I = \left(\prod_k \int d\xi_k\right) \exp\left(-\xi_i B_{ij} \xi_j\right) \tag{240}$$

where B is a symmetric matrix with eigenvalues b_i. Diagonalize B via orthogonal matrix $\mathscr{O}$, and switch to x_i variables defined by

$$\xi_i = \mathscr{O}_{ij} x_j \tag{241}$$

Then

$$\begin{aligned} I &= \left(\prod_k \int d\xi_k\right) \exp\left(-\xi_i B_{ij} \xi_j\right) \\ &= \left(\prod_k \int d\xi_k\right) \exp\left(-\sum_i b_i x_i^2\right) \\ &= \prod_i \left(\int dx_i \exp\left(-\sum_i b_i x_i^2\right)\right) \\ &= \prod_i \sqrt{\frac{\pi}{b_i}} \\ &= \text{const.} \times \frac{1}{\sqrt{\det(B)}} \end{aligned}$$

In other words,

$$\left(\prod_k \int d\xi_k\right) e^{-\xi_i B_{ij} \xi_j} = (\text{const.})(\det(B))^{-1/2} \tag{242}$$

We will see later on that formulae like this one will arise in multiple situations in cases with quantum fields of spins higher than zero, except that the power of the determinant is different. In particular, we will see soonish that fermionic fields give rise to a determinant to a positive power.

Now let us stare at what we just computed. Notice that

$$S_0[\phi] = \tfrac{1}{2} \int d^D x \left[\phi\left(-\partial^2 - m^2\right)\phi\right] + (\text{surface term}) \tag{243}$$

So, provided integration by parts makes sense in our spacetime and with our field theory, formally we can write

$$B = \left(-\partial^2 - m^2 \mathbb{1}\right) \tag{244}$$

as a matrix. Therefore

$$Z_\phi = \int \mathscr{D}\phi \exp\left(iS_0[\phi]\right) = (\text{const}) \cdot \left[\det(m^2 + \partial^2)\right]^{-1/2} \tag{245}$$

This thing is known as a functional determinant, as mentioned earlier. It may seem ill-defined, but because of what we learned in the previous subsection we know that in fact any wildly oscillating phases in the expression for Z_ϕ will cancel out of expressions for physical correlation functions.

Let us now see how the Wick contraction formula familiar from canonical quantization makes an appearance in the functional path integral approach, starting with the two-point correlation function. To see how that comes about we need to Fourier expand our fields living on the square spacetime lattice:

$$\phi(x_1) = \frac{1}{V}\sum_m e^{-ik_m\cdot x_1}\phi_m \tag{246}$$

Then, in computing the time-ordered product (233) appropriate for the two-point correlation function, we have for the numerator

$$\begin{aligned}\text{numerator} \quad &= \quad \left(\prod_{n|k_n^0>0}\int d(\Re\phi_n)d(\Im\phi_n)\right)\frac{1}{V^2}\sum_{m,\ell} e^{-i(k_m\cdot x_1+k_\ell\cdot x_2)}\times \\ &\qquad \times(\Re\phi_m+i\Im\phi_m)(\Re\phi_\ell+i\Im\phi_\ell)\exp\left(-\frac{i}{V}\sum_{n|k_n^0>0}(m^2-k_n^2)\left[(\Re\phi_n)^2+(\Im\phi_n)^2\right]\right)\end{aligned}$$

Many of the terms in this expression are zero by symmetry. In fact, the only surviving terms arise from when

$$k_m = \pm k_\ell \tag{247}$$

Suppose for definiteness that $k_n^0 > 0$. Inspecting $k_m = +k_\ell$, we can see that the term involving $(\Re\phi_n)^2$ is nonzero but is exactly cancelled by the term involving $(\Im\phi_n)^2$. For $k_m = -k_\ell$, because of the reality condition on the scalar field an extra relative $-$ sign arises between the $(\Re\phi_n)^2$ and $(\Im\phi_n)^2$ pieces, and so they add rather than cancelling. If instead $k_n^0 < 0$, the final expressions end up being very similar.

Doing the integrals over $\Re\phi_n$ and $\Im\phi_n$ gives for the two-point correlation function numerator

$$\text{numerator} = \frac{1}{V^2}\sum_m e^{-ik_m(x_1-x_2)}\left[\prod_{k_n^0>0}\frac{-i\pi V}{(m^2-k_n^2)}\right]\left[\prod_{k_m^0>0}\frac{-i\pi V}{(m^2-k_m^2)}\right] \tag{248}$$

Let's inspect this expression carefully. If we had not included the two field operators in the numerator (in order to find the two point correlator), then the first [...] factor in the above equation would have been already obtained. The new part, the bit with the meat in it, is the second factor. This is in fact none other than the *Feynman propagator* in discretized form. So in the continuum limit,

$$\langle 0|T\left\{\phi(x_1)\phi(x_2)\right\}|0\rangle = \int\frac{d^Dk}{(2\pi)^D}\frac{ie^{-ik\cdot(x_1-x_2)}}{(k^2-m^2+i\epsilon)} = D_F(x_1-x_2) \tag{249}$$

This is *exactly* the same as was obtained in QFT1 from the canonical quantization approach.

How about higher correlation functions? Well, one thing we can see immediately by symmetry from the properties of Gaussian integrals is that in free scalar field theory

$$\langle 0|T\left\{\phi(x_1)\phi(x_2)\phi(x_3)\right\}|0\rangle \equiv 0 \tag{250}$$

The next least trivial task is to compute the four-point correlator. Again, we will display how to compute it here in gory detail, but in future we won't repeat the same level of detail.

For the four-point correlator, by analogy with the two-point case we'll be looking at a numerator with the Fourier expansions of the four field operators $\phi = (\Re\phi + \Im\phi)$ at positions (x_1, x_2, x_3, x_4) with Fourier mode numbers labelled by (m, ℓ, p, q) respectively. Again, by symmetry, many of the terms in the integrand are zero. Nonvanishing pieces will show up, as with the two-point correlator case, whenever particular *pairs* of ϕs pair up rather than cancelling out.

Consider one of these nonzero contributions: it occurs when

$$k_\ell = -k_m, \qquad k_q = -k_p \tag{251}$$

After performing the Gaussian integrals, we find

$$\begin{aligned}
&\frac{1}{V^4}\sum_{m,p} e^{-ik_m\cdot(x_1-x_2)} e^{-ik_p\cdot(x_3-x_4)} \left(\prod_{n|k_n^0>0} \frac{-i\pi V}{(m^2-k_n^2)}\right) \times \\
&\times \left(\prod_{n|k_n^0>0} \frac{-i\pi V}{(m^2-k_m^2+i\epsilon)}\right)\left(\prod_{n|k_n^0>0} \frac{-i\pi V}{(m^2-k_p^2+i\epsilon)}\right) \\
&\underset{V\to\infty}{\longrightarrow} \left(\prod_{n|k_n^0>0} \frac{-i\pi V}{(m^2-k_n^2)} D_F(x_1-x_2)D_F(x_3-x_4)\right)
\end{aligned}$$

Therefore, one part of the four-point correlation function is $D_F(x_1-x_2)D_F(x_3-x_4)$. This should look familiar! :D It is just our old friend the Wick Contraction, applied between the first and third field operators in the time-ordered product. By looking at the other pairings of momenta k_ℓ, k_m, k_q, k_p, we can see that there are two other terms which are identical except for a shuffling of indices. Putting all the pieces together and summing (over all full contractions, in the language of canonical quantization) gives the four-point function

$$\begin{aligned}
\langle 0|T\left\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_4)\right\}|0\rangle \;=\; & D_F(x_1-x_2)D_F(x_3-x_4)+ \\
& D_F(x_1-x_3)D_F(x_2-x_4)+ \\
& D_F(x_1-x_4)D_F(x_2-x_3)
\end{aligned}$$

Notice that the quantity $:\mathscr{T}\exp(-i/\hbar\int dt\hat{H}):$ appears nowhere in this analysis. Nonetheless, our newfangled functional approach produced the same answer as the canonical method.

How about general potentials? Can we say anything when the field theory is not free?

The Feynman path integral approach is of course well-suited to general potentials, including those higher than quadratic in fields. The main restriction, as with the canonical approach, is that perturbation theory is the main vector for analysing the physics. Treating the potential as a small perturbation about the free action, we perform an expansion

$$\exp\left(\frac{i}{\hbar}\int d^D x \mathscr{L}\right) = \exp\left(\frac{i}{\hbar}\int d^D x \mathscr{L}_{\text{free}}\right) \times \left[1 + \frac{i}{\hbar}\int d^D x \mathscr{L}_{\text{int}}(\phi) + \cdots\right] \tag{252}$$

For instance, for

$$L_{\text{int}} = -\frac{\lambda}{k!}\phi^k \tag{253}$$

for small λ we can write

$$e^{iS[\phi]/\hbar} = e^{iS_0/\hbar}\left(1 - \frac{i}{\hbar}\int d^D x \frac{\lambda}{k!}\phi^k + \cdots\right) \tag{254}$$

Notice that

$$\frac{i}{\hbar}\int d^D x \mathscr{L}_{\text{int}} = -\frac{i}{\hbar}\int d^D x \mathscr{V}_{\text{int}} = -\frac{i}{\hbar}\int d^D x \mathscr{H}_{\text{int}} \tag{255}$$

which starts to smell an awful lot like canonical quantization. In fact, as we will see, perturbation theory ends up being *identical.* More explicitly, by using our generalized formula (233)

$$\langle\Omega|T\left\{\phi_H(x_1)\cdots\phi_H(x_n)\right\}|\Omega\rangle = \lim_{T\to\infty(1-i\epsilon)} \frac{\int \mathscr{D}\phi \cdot \phi(x_1)\cdots\phi(x_n)\cdot \exp\left(i\int_{-T}^{+T}\mathscr{L}\right)}{\int \mathscr{D}\phi \exp\left(i\int_{-T}^{+T}\mathscr{L}\right)} \tag{256}$$

It is straightforward, if tedious, to see that again

- The prefactor piece
$$\prod_{n|k_n>0}\frac{-i\pi V}{(m^2-k^2)} \tag{257}$$
cancels between the numerator and denominator.
- The full interacting correlation function is expressed solely in terms of *free* correlators.
- The combinatorics work exactly as in canonical quantization. In particular, all disconnected bubble diagrams exponentiate.

The reason we won't be going into gory detail to compute this the same way we computed the two-point correlation function earlier is that there's a *much* slicker way of doing it – by taking advantage of the power of functional calculus. Instead of dealing with all these awkward Fourier transforms, integrations, etc, we'll introduce the concept of a *source* in the action. This single clever concept

$$S \to S + S_{\text{source}} = S + \int \phi \cdot J[\phi] \tag{258}$$

will allow us to generate correlation functions via functional differentiation of the Feynman path integral with respect to the source $J(\phi)$. It also bears a remarkable resemblance to a field Legendre transformation, which is no accident.

5 Generating functionals

First, a lightning review of the salient features of functional differentiation. The axioms for functional differentiation are

$$\frac{\delta}{\delta J(x)}\int d^D y J(y)\phi(y) = \phi(x)\,, \qquad \text{or } \frac{\delta}{\delta J(x)} J(y) = \delta^{(D)}(x-y) \tag{259}$$

For example,

$$\frac{\delta}{\delta J(x)} \exp\left(i\int d^D y J(y)\phi(y)\right) = i\phi(x)\exp\left(i\int d^D y J(y)\phi(y)\right) \tag{260}$$

Integration by parts can prove handy. For instance:

$$\frac{\delta}{\delta J(x)}\int d^D y\,(\partial_\mu J(y))\,V^\mu(y) = -\partial_\mu V^\mu(x) \tag{261}$$

5.1 The generating functional for all Feynman graphs

We now have enough background to define the generating functional $Z[J]$[9]:

$$Z_{\text{scalar}}[J] \equiv \int \mathscr{D}\phi \exp\left[\frac{i}{\hbar}\int d^D x\,\{\mathscr{L} + J(x)\phi(x)\}\right] \tag{262}$$

We will find correlation functions by operating on $Z[J]$ with functional derivatives $\delta/\delta J$.

The general n-point correlation function is then

$$\langle 0|T\{\phi(x_1)\cdots\phi(x_n)\}|0\rangle = \frac{1}{Z}\left(-i\frac{\delta}{\delta J(x_1)}\right)\cdots\left(-i\frac{\delta}{\delta J(x_n)}\right)Z[J]\Big|_{J=0} \tag{263}$$

This formula holds for interacting scalar field theories. For *free* scalar field theory it can be rearranged in a very explicit way which makes extracting correlation functions a piece of cake compared to how we computed them earlier, so we now show how this works.

Consider

$$\int d^D x\,(\mathscr{L}_0 + J\cdot\phi) = \int d^D x\left[\tfrac{1}{2}\phi\left(-\partial^2 - m^2 + i\epsilon\right)\phi + J\cdot\phi\right] \tag{264}$$

Let's complete the square here by defining

$$\phi'(x) \equiv \phi(x) - i\int d^D y \mathscr{D}_F(x-y)J(y) \tag{265}$$

where $\mathscr{D}_F(x-y)$ is our friend the Feynman propagator. In terms of this shifted field,

$$\int d^D x\,(\mathscr{L}_0 + J\cdot\phi) \;=\; \int d^D x\left[\tfrac{1}{2}\phi'\left(-\partial^2 - m^2 + i\epsilon\right)\phi'\right]$$

[9]To avoid confusion, we will use the subscript 0 to refer to a free field theory, not a spin zero field theory.

$$-\int d^D x \int d^D y \tfrac{1}{2} J(x) \left[-i\mathscr{D}_F(x,y)\right] J(y)$$

More formally, we can write

$$\phi' \equiv \phi + \left(-\partial^2 - m^2 + i\epsilon\right)^{-1} J \tag{266}$$

because $\mathscr{D}_F$ is (as you will know from PHY2403F) the Green's Function of the Klein-Gordon operator.

Our expression becomes

$$\int d^D x \left[\mathscr{L}_0 + J\cdot\phi\right] = \int d^D x \left[\tfrac{1}{2}\phi' \left(-\partial^2 - m^2 + i\epsilon\right)\phi' - \tfrac{1}{2} J \left(-\partial^2 - m^2 + i\epsilon\right)^{-1} J\right] \tag{267}$$

From a functional perspective, our shifted field is literally just a shift: the quantity

$$\phi' = \phi - i \int d^D y \mathscr{D}_F(x-y) J(y) \tag{268}$$

does not involve ϕ explicitly and so the Jacobian of the transformation is unity.

$$\mathscr{D}\phi' = \mathscr{D}\phi \tag{269}$$

Therefore,

$$Z[J] = \int \mathscr{D}\phi' \exp\left(i\int d^D x \mathscr{L}_0(\phi')\right) \exp\left(-i\int d^D x d^D y \tfrac{1}{2} J(x) \left[-i\mathscr{D}_F(x-y)J(y)\right]\right) \tag{270}$$

Notice that the first factor here is simply the free path integral. The other piece is *independent* of ϕ', so we can write

$$Z_{\rm KG}[J] = Z_0 \exp\left(-\tfrac{1}{2}\int d^D x d^D y J(x)\mathscr{D}_F(x-y)J(y)\right) \tag{271}$$

This little devil of a formula is extremely useful, as we can now see explicitly by using it to compute the free two- and four-point functions that we so laboriously extracted earlier by discretizing spacetime and tussling with Fourier expansions.

For the two-point function:

$$\begin{aligned}
& \langle 0| T\left\{\phi(x_1)\phi(x_2)\right\}|0\rangle \\
= & -\frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_1)} \exp\left[-\int d^D x d^D y \tfrac{1}{2} J(x)\mathscr{D}_F(x-y)J(y)\right]\Bigg|_{J=0} \\
= & -\frac{\delta}{\delta J(x_1)}\left\{\left[-\tfrac{1}{2}\int d^D y \mathscr{D}_F(x_2-y)J(y) - \tfrac{1}{2}\int d^D x J(x)\mathscr{D}_F(x-x_2)\right]\frac{Z[J]}{Z_0}\right\}\Bigg|_{J=0} \\
= & \mathscr{D}_F(x_1-x_2)
\end{aligned}$$

Notice that taking one $\delta/\delta J$ brought down two identical factors, each of which had a J in it. Therefore, the second $\delta/\delta J$ only gave terms surviving at $J=0$ when it acted on the $\{\cdots\}$ factor, not on the $Z[J]/Z_0$ part.

Let's do the four-point function too:

$$
\begin{aligned}
& \langle 0|T\{\phi(x_1)\phi(x_2)\phi(x_3)\phi(x_3)\}|0\rangle \\
= & \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_4)}\exp\left(\tfrac{1}{2}\int d^Dx d^Dy J(x)\mathscr{D}_F(x-y)J(y)\right)\Bigg|_{J=0} \\
= & \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_3)}\left(-J(x)\mathscr{D}_F(x-x_4)\right)\exp\left(-\tfrac{1}{2}\int d^Dx d^Dy J(x)\mathscr{D}_F(x-y)J(y)\right)\Bigg|_{J=0} \\
= & \frac{\delta}{\delta J(x_1)}\frac{\delta}{\delta J(x_2)}\left[\mathscr{D}_F(x_3-x_4)+J(x)\mathscr{D}_F(x-x_4)J(y)\mathscr{D}_F(y-x_3)\right]e^{-\frac{1}{2}J\mathscr{D}_F J}\Bigg|_{J=0} \\
= & \frac{\delta}{\delta J(x_1)}\left[\mathscr{D}_F(x_3-x_4)J(x)\mathscr{D}_F(x-x_2)\right. \\
& \quad \left.+\mathscr{D}_F(x_2-x_4)J(y)\mathscr{D}_F(y-x_3)+J(x)\mathscr{D}_F(x-x_4)\mathscr{D}_F(x_2-x_3)\right]e^{-\frac{1}{2}J\mathscr{D}_F J}\Bigg|_{J=0} \\
= & \mathscr{D}_F(x_3-x_4)\mathscr{D}_F(x_1-x_2)+\mathscr{D}_F(x_2-x_4)\mathscr{D}_F(x_1-x_3)+\mathscr{D}_F(x_1-x_4)\mathscr{D}_F(x_2-x_3)
\end{aligned}
$$

Yay! That was easy! :D

In summary: by using the source trick

$$S \longrightarrow S + \int d^Dx\ \phi \cdot J[\phi] \tag{272}$$

we found the two-point and four-point correlation functions for the free scalar field theory. Moreover, the functional differentiation process we went through straightforwardly generalizes to give the n-point correlation functions for free scalar field theory. The resulting expression is:

$$\frac{\int [d\phi] e^{iS_0[\phi]}\phi(y_1)\cdots\phi(y_n)}{\int [d\phi] e^{iS_0}} = \sum_{\text{pairs}(ij)}\prod \mathscr{D}_F(y_i-y_j) \tag{273}$$

This is of course just Wick's Theorem in disguise. Here, as distinct from PHY2403F, we derived it using functional calculus: **no field operators were harmed in the production of this formula.**

5.2 Analogy between statistical mechanics and QFT

Functionally differentiating w.r.t. a source term linear in $J(x)$ in the Feynman phase $e^{iS/\hbar}$ is similar in spirit to the mechanism we often use to compute thermal average behaviour in statistical mechanical systems. To pin this down more precisely, let us think carefully about contours in the complex momentum plane. Recall that we had the prescription

$$\frac{1}{(k^2-m^2)} \longrightarrow \frac{1}{(k^2-m^2+i\epsilon)} \tag{274}$$

in propagators, in order to render the Feynman path integral well-defined. This $+i\epsilon$ prescription 'tips' the contour into the complex plane in 'under-and-over' style just the right direction to allow rotation *counterclockwise* onto the imaginary axis.

This motivates us to wonder what the Wick rotated version of the Feynman path integral (FPI) might be physically. Let us rotate

$$t = -i\tau \tag{275}$$

Then

$$x \cdot x = t^2 - |\vec{x}|^2 = -\tau^2 - |\vec{x}|^2 \equiv -x_E^2 \tag{276}$$

By looking at each Feynman diagram in turn, it can be shown that the analytic continuation of the time variables in any Green's function of a QFT gives a correlation function invariant under rotational symmetry of D-dimensional Euclidean space. In D dimensions this symmetry group is $SO(D)$ rather than the Lorentz group $SO(1, d = D - 1)$.

To end this exposition of functional quantization for scalar field theory, let's now do a specific example to see the Stat Mech $\leftrightarrow$ QFT connection. Suppose we pick ϕ^4 theory for definiteness. The Lorentzian signature action reads

$$S_L = \int d^D x (\mathscr{L} + J \cdot \phi) = \int d^D x \left[\tfrac{1}{2} \left\{ (\partial\phi)^2 - m^2\phi^2 \right\} - \frac{\lambda}{4!}\phi^4 + J \cdot \phi \right] \tag{277}$$

Wick rotating gives

$$S_L = -i \int d^D x_E \left[\tfrac{1}{2} \left\{ -(\partial_E\phi)^2 - m^2\phi^2 \right\} - \frac{\lambda}{4!}\phi^4 + J \cdot \phi \right] = i \int d^D x_E (\mathscr{L}_E - J \cdot \phi) \tag{278}$$

Since $e^{iS_L/\hbar} = e^{-S_E/\hbar}$, this RHS of the above equation is just the expression for the Helmholtz free energy of a D-dimensional ferromagnet in the Landau theory in disguise, if $\phi(x_E)$ plays the role of the fluctuating spin field, and J plays the role of an external magnetic field.

The Wick rotated generating functional becomes

$$Z[J] = \int \mathscr{D}\phi \exp\left(- \int d^D x_E \left[\mathscr{L}_E - J \cdot \phi \right] \right) \tag{279}$$

Note that this is well-definined – when the Euclidean action is large this happens either because the field or field gradients get large. No oscillatory behaviour here; just damping. The groundstate energy is bounded from below.

The exponential that we see above is therefore exactly like a statistical weighting factor (recall the grand canonical partition function). Inspecting the two-point function for $\lambda = 0$, in direct analogy to before we find

$$\langle 0 | T \left\{ \phi(x_{E,1})\phi(x_{E,2}) \right\} | 0 \rangle = \int \frac{d^D k_E}{(2\pi)^D} \frac{\exp\left(i k_E \cdot (x_{E,1} - x_{E,2}) \right)}{(k_E^2 + m^2)} \tag{280}$$

In other words, this is the Feynman propagator in the *spacelike* region. For massive fields, the coordinate space Yukawa potential falls off as $e^{-m|x_{E,1} - x_{E,2}|}$. Therefore, the Compton wavelength $\lambda = h/(mc)$ of quanta in QFT translates in Stat Mech language to the correlation length for statistical fluctuations. For more cute observations about crossovers between theoretical HEP and theoretical condensed matter, try the Quantum Field Theory in a Nutshell book by Anthony Zee.

5.3 Feynman rules for scalar field theory

In position space, the Feynman rules for spin-zero field theory can be summarized as follows.

- Write the potential as $V(\phi) = g\sum_j v_j \phi^j/j!$
- At $\mathscr{O}(g^n)$, contributions like $v_{j_1}\cdots v_{j_n}$ will arise; associate this with a diagram with n vertices. The ith vertex has j_i lines emanating from it.
- Each propagator is denoted by a line, and must either connect two vertices (denoted by ●——●) or connect a vertex to an external source (denoted by ●——×)
- The E-point Green's functions have E external lines. The number of internal lines I is given by $2I = \sum_i j_i - E$.
- Each internal line connecting two points (x_1, x_2) is represented by a Feynman propagator $\mathscr{D}_F(x_1 - x_2)$.
- Each external line gets a factor $\mathscr{D}_F(y_i - x_a)J(x_a)$.
- Integrate over internal and external points. For just an n-point Green's function, omit the integrals over external points and $J(x_a)$ pieces.
- A Feynman graph at nth order comes with a factor $(ig)^n v_{j_1}\cdots v_{j_n} S_G$ where S_G is a combinatoric factor, equal to the number of times a graph can be obtained in a sum over all possible pairings (normalized by the $1/j!$ bits in the definition of v_{j_i}).

What changes if we switch to momentum space?

- The Feynman propagator becomes: $\mathscr{D}_F(k) = i/(k^2 - m^2 + i\epsilon)$.
- Instead of integrating over positions of points, integrate over loop momenta $\int d^Dk/(2\pi)^D$.
- Enforce momentum conservation at each vertex with incoming momenta $\{p_i\}$ via delta functions $(2\pi)^D\delta^D(\sum_i p_i)$.
- Enforce overall momentum conservation.

This list of steps allows us to compute the physics of the interacting scalar field theory, in the approximation $|g| \ll 1$ where perturbation theory is valid.

5.4 Generating functional for connected graphs $W[J]$

Previously, we calculated the generating functional for free scalar field theory – as a prélude to doing the calculation for interacting field theories. We found

$$Z_0[J] = \exp\left(-\tfrac{1}{2}\int J(x)\mathscr{D}_F(x-y)J(y)\right) \tag{281}$$

Let us now use this knowledge to motivate definition of another generating functional $W[J]$, which is closely related to $Z[J]$ but different in important ways. The physical utility of the new generating functional $W[J]$ is that it allows us to focus only on connected Feynman graphs, which are the ones of interest for computing physical scattering amplitudes.

Denoting by N the normalization constant in the FPI, we had

$$\frac{Z_0[J]}{N} = 1 - \frac{1}{2}\int J\mathscr{D}_F J + \frac{1}{2!}\left(\frac{1}{2}\int J\mathscr{D}_F J\right)^2 - \frac{1}{3!}\left(\frac{1}{2}\int J\mathscr{D}_F J\right)^3 \cdots$$

$$= \quad 1 - \frac{1}{2} \times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times + \frac{1}{2!}\left(\frac{1}{2}\right)^2 \begin{matrix}\times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times \\ \times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times\end{matrix} \; - \frac{1}{3!}\left(\frac{1}{2}\right)^3 \begin{matrix}\times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times \\ \times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times \\ \times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times\end{matrix} \tag{282}$$

Look at the structure we are seeing here. The physics is driven solely by $\times\!\!\!-\!\!\!-\!\!\!-\!\!\!\times$. Mathematically, because of the combinatorics it is manifest that this exponentiates to give $Z_0[J]$. Therefore, we define a new $W[J]$ by

$$Z[J] \equiv \exp\left(iW[J]\right) \tag{283}$$

Of course, for the free scalar field theory example, we would have $W_0[J] = -\int \frac{1}{2} J \mathscr{D}_F J$.

The really amazing thing about $W[J]$ is that it has contributions from connected Feynman graphs only. We have seen this directly above in the free case, but it is also pretty straightforward to show that this conclusion holds not only for free field theories but also for interacting QFTs as well (at least in perturbation theory). The underlying reason is the structure of the Feynman rules, as written above.

The major physics concept introduced here is that **the sum over all Feynman graphs is the exponential of the sum over connected Feynman graphs**. This follows in the interacting case from (a) the Feynman rules and (b) the combinatorics of the symmetry-factor S_G, because a graph G with k identical disconnected pieces has an obvious $\mathscr{S}_k$ symmetry with order $k!$. As we develop explicit examples at one-loop level we will see this illustrated in more detail.

One awesome thing about $W[J]$ is that the Feynman rules for $Z[J]$ and $W[J]$ are identical, except for $W[J]$'s omission of disconnected graphs. Such disconnected graphs will cancel out of physical correlation functions anyway.

5.5 Counting powers of the coupling constant

Suppose that by a clever rescaling of fields we can write

$$S[\phi] = \frac{1}{g^2} s(\phi) \tag{284}$$

where $|g| \ll 1$ is a small dimensionless parameter and $s[\phi]$ contains only nonnegative powers of g. In such a physical situation we can immediately deploy the stationary phase approximation, a.k.a. the method of steepest descents. This focuses us on stationary points where

$$\frac{\delta s}{\delta \phi(x)} + J(x) = 0 \tag{285}$$

i.e., the classical field equation in the presence of the source J.

Define a new field - a hybridized function/functional beastie - which is a regular function of x while being a *functional* of J. This $\phi(x, [J])$ is the classical solution to the above stationary phase equation. The connection between $\phi(x, J[x])$ and the connected generating functional $W[J]$ is

$$W_{\rm cl}[J] = \frac{1}{g^2}\left(s[\phi[J]] + \int J\phi[J]\right) \tag{286}$$

In the classical approximation, then, the connected generating functional is the Legendre transformation of the classical action. Incidentally, solving the classical field equations with a source J of $\mathscr{O}(1)$ sums up all tree diagrams, with any number of external legs.

If the functional $s[\phi]$ is g-independent, then the semiclassical expansion also has a topological interpretation in terms of Feynman graphs. To see this, start by denoting the stationary point of the $J=0$ functional integral as ϕ_0. Then the first correction to the classical action is

$$S_{(2)}[\phi_0, \delta\phi] \equiv \frac{1}{g^2}\int d^Dx d^Dy \delta\phi(x)\delta\phi(y)\frac{\delta^2 s}{\delta\phi(x)\delta\phi(y)}\bigg|_{\phi=\phi_0} \tag{287}$$

Define a rescaled field fluctuation by

$$\Delta \equiv g\delta\phi \tag{288}$$

Then $\mathscr{O}(1)$ fluctuations in Δ correspond to $\mathscr{O}(g)$ fluctuations in our original fields $\delta\phi$, and

$$\frac{1}{g^2}s[\phi] = \frac{1}{g^2}s[\phi_0] + s_2[\Delta] + \sum_j g^{j-2}v_j[\Delta] \tag{289}$$

To work out which terms contribute to an E-point function at $\mathscr{O}(g^m)$, consider our vertices v_{j_i} in a Feynman graph.

- The number of lines $V = \sum_i j_i$ emanating from a vertex has to sum up to n:

$$\sum_i (j_i - 2) = \sum_i j_i - 2V = n \tag{290}$$

  Note that the -2 in the $(j_i - 2)$ term is present because of the scaling (289) above.
- The number of internal lines I is

$$I = \tfrac{1}{2}\sum_i j_i - E \tag{291}$$

- The number of loops for a connected diagram is

$$L = I - V + 1 = \tfrac{1}{2}\sum_i j_i - E - V + 1 = n/2 - E + 1 \tag{292}$$

Therefore, for a fixed number of external lines the g^2 expansion *is* the loop expansion. This is what we set out to show.

5.6 Quantum [effective] action and 1PI diagrams

Before we introduce $\Gamma[\phi]$, let us do a lightning review of usefulness of Legendre transformations. Consider a curve $f(x)$. What equation does the *tangent* to the curve possess? Let g be the intercept on the f axis and let $u = df/dx$ be the gradient. Then the tangent to the curve has equation

$$f = ux + g\,.$$

Now suppose that we wish to pick u as the independent variable, and let g be the dependent one, i.e., $g = f - ux$. How do we reconstruct $f(x)$, if we know only $g(u)$? The problem is that knowing $g(u)$ does not uniquely reconstruct f; it fails to do so because derivatives do not see constant shifts. So we use the **envelope of tangents** method,

$$g(u) = f(x) - ux\,.$$

Let us recall a thermodynamics example. Suppose we have $U = U(S,V)$. But S, being extensive, does not lend itself to measurement nearly as easily as temperature does! So let us define $F(T,V) = U(S,V) - TS$, where

$$T = \left(\frac{\partial U}{\partial S}\right)_V$$

which follows from the first law, $dU = TdS - pdV$. The analogy here between path integrals and statistical methanics is actually very deep:-

Feynman path integral	partition function
$Z_{FPI} = e^{iW}$	$Z_{SM} = e^{-\beta F}$
$W[J] = \Gamma[\phi] + \int J\phi$	$F(T) = U(S) - TS$

Based on this knowledge, let us now introduce a Legendre-transformed connected generating functional $\Gamma[\phi[J]]$ via

$$W[J] \equiv \Gamma[\phi[J]] + \int J\cdot\phi[J] \tag{293}$$

Then $\phi(x,[J])$ is a solution of

$$\frac{\delta W}{\delta J(x)} = \phi(x,[J]) \tag{294}$$

The quantity $\Gamma[\phi(x,[J])]$ is called the quantum action, sometimes confusingly called the quantum effective action. The expansion coefficients of Γ

$$\Gamma[\phi] = \sum_n \frac{1}{n!}\int d^Dx_1\cdots d^Dx_n \Gamma_n(x_1\cdots x_n)\phi(x_1)\cdots\phi(x_n) \tag{295}$$

are called **one-particle irreducible** (1PI) Green's functions. Diagramatically, we construct 1PI Green's functions as the sum of all graphs that cannot be cut into disconnected parts by chopping one propagator line. Note that, at tree level, vertices are straightforwardly obtained from the classical action.

You should think of 1PI diagrams as being a bit like Lego blocks. To give the punch line first: connected Green's functions are built from tree diagrams with

- 1PI functions as vertices;
- propagators as full connected two-point functions.

How would we prove this assertion? Begin with

$$W[J] = \Gamma[\phi[J]] + \int J\cdot\phi[J] \tag{296}$$

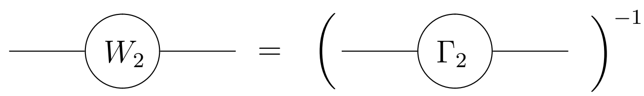
Since

$$\frac{\delta^2 W}{\delta J(x)\delta J(y)} = \frac{\delta}{\delta J(x)}\left[\phi(y)\right] = \left(\frac{\delta\phi(x)}{\delta J(y)}\right) = \left(\frac{\delta J(y)}{\delta\phi(x)}\right)^{-1} = -\left(\frac{\delta^2\Gamma}{\delta\phi(x)\delta\phi(y)}\right)^{-1} \tag{297}$$

Note that this inverse is a functional inverse, in the sense of

$$\int d^D x K(x,z)K^{-1}(z,y) = \delta^D(x-y) \tag{298}$$

for any K. In discretized spacetime, this is matrix inversion and (297) is the continuum limit of it. Another way to write (297) is graphically:



We can derive more by differentiating w.r.t. J:

$$\frac{\delta^3 W}{\delta J(x)\delta J(y)\delta J(z)} = -\frac{\delta}{\delta J(z)}\left(\Gamma_2^{-1}(x,y)\right) \tag{299}$$

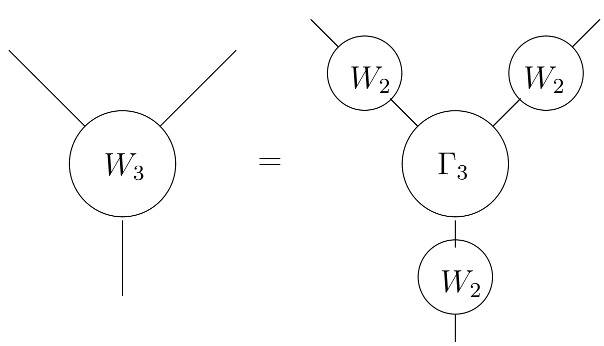
Our next necessity here is to define how to differentiate the functional inverse. It is simplest to use the discretization of spacetime regularization, and use matrix discretization as a guide. We know that for invertible matrices K, $d(K^{-1}) = -K^{-1}dKK^{-1}$, and so by analogy

$$W_3(x,y,z) = \int d^D w_1 \int d^D w_2 W_2(x,w_1)W_2(y,w_2)\frac{\delta\Gamma_2(w_1,w_2)}{\delta J(z)} \tag{300}$$

Now, by definition, $W[J] = \Gamma[\phi[J]] + \int J\cdot\phi[J]$, and so $\phi = \delta W[J]/\delta J$. Using that and the chain rule for differentiation gives

$$W_3(x,y,z) = \int d^D w_1 \int d^D w_2 \int d^D w_3 W_2(x,w_1)W_2(y,w_2)W_2(z,w_3)\Gamma_3(w_1,w_2,w_3) \tag{301}$$

Graphically, this is



You will be working out the $n=4$ case as part of the second homework assignment, HW2.

5.7 The generating functional and the Schwinger-Dyson equations

Let us now turn to connecting our powerful newfangled functional methods for doing quantum field theory back to our knowledge from canonical quantization.

The story of the Schwinger-Dyson equations relies on two familiar pieces of physics from canonical quantization: (a) the canonical commutation relations (CCRs) and (b) the Heisenberg equations of motion for the field operators. Their basic idea was to derive a closed set of equations for *all* n-point Green's functions. The resulting Schwinger-Dyson equations give $Z[J]$ as that generating functional, and are practically useful as well. Note that in this subsection ϕ denotes the field *operator*. Suppose, for definiteness, that we pick the kinetic term to be quadratic in field derivatives:

$$\mathscr{L} = \tfrac{1}{2}(\partial\phi)^2 - V(\phi) \tag{302}$$

where $V(\phi)$ may include a mass term of the form $\frac{1}{2}m^2\phi^2$.

Since n-point Green's functions are

$$\langle 0|T\left\{\phi(x_1)\cdots\phi(x_n)\right\}|0\rangle \tag{303}$$

it follows immediately that the *generating functional* for these Green's functions (in the mathematical sense) is none other than the vacuum persistence amplitude in the presence of a source J:

$$Z[J] \equiv \langle 0|T\left\{\exp\left(\frac{i}{\hbar}\int d^Dx\, J(x)\cdot\phi(x)\right)\right\}|0\rangle \tag{304}$$

Note that this equation is not the same expression as we had for the FPI: it is a canonical beastie.

We have to be careful to keep track of the effect of time-ordering. For instance, consider

$$\begin{aligned}
&\partial_0^2\langle 0|T\left\{\phi(x)\phi(y)\right\}|0\rangle \\
&= \partial_0\left[\langle 0|\left(T\left\{\partial_0\phi(x)\phi(y) + \delta(x^0-y^0)[\phi(x),\phi(y)]\right\}\right)|0\rangle\right] \\
&= \langle 0|T\left\{\partial_0^2\phi(x)\phi(y)\right\}|0\rangle - i\delta^D(x-y)
\end{aligned} \tag{305}$$

Here we used the CCRs $[\phi(x),\Pi(y)] = i$, where Π is the field canonical momentum. Also, notice something important about the intermediate step above. Since field operators $\phi(x)$ commute at equal times, it might have appeared that the second term could be dropped. That presumption would be incorrect, because it gives the wrong answer when further derivatives are taken: it misses the contact term.

As an alert reader can check explicitly, the above expression can be generalized. The result is:

$$\begin{aligned}
&\partial_0^2\langle 0|T\left\{\phi(x)\phi(y_1)\cdots\phi(y_n)\right\}|0\rangle \\
&= \langle 0|T\left\{\partial_0^2\phi(x)\phi(y_1)\cdots\phi(y_n)\right\}|0\rangle \\
&\quad -i\sum_j \delta^D(x-y_j)\langle 0|T\left\{\phi(y_1)\cdots\phi(y_{j-1})\phi(y_{j+1})\cdots\phi(y_n)\right\}|0\rangle
\end{aligned} \tag{306}$$

This expression, involving subtle delta function pieces, is the nugget of the Schwinger-Dyson story. The rest of our calculation will amount to using this fact that **correlation functions obey differential equations involving contact terms (bits with delta functions)**.

Next, we multiply through by $J(y_1)\cdots J(y_n)$ and integrate against the nugget. For our next trick, we then recruit the Heisenberg equations of motion for $\partial_0\phi$ to substitute for time derivatives of field operators. You should check explicitly on your own that the resulting equation is

$$\partial^2\left(-i\frac{\delta Z}{\delta J(x)}\right) = -\left\{\frac{\partial V}{\partial\phi}\left[-i\frac{\delta}{\delta J(x)}\right] + J(x)\right\} Z[J] \tag{307}$$

The partial derivatives ∂ in this expression are fully relativistic, and $Z[J]$ is the generating functional for all [canonical] Green's functions.

Our next step is most convincing if we work in a regularization where spacetime is discretized. We write the nth Fourier mode of $\phi(x)$ as ϕ_n and the nth Fourier mode of $J(x)$ as J_n. In this discretized context, ∂^2 becomes a (symmetric) matrix K_{mn}. In discretized form, then, the Schwinger-Dyson equations become

$$K_{mn}\cdot -i\frac{\partial Z}{\partial J_n} = -\left[V'\left(-i\frac{\partial}{\partial J_m}\right) + J_m\right] Z[J] \tag{308}$$

Now for the crucial point. Notice that the only *explicit* J-dependence in this expression is linear. Since for Fourier transforms $i\partial$ acts like $-ik$, the Fourier transform of the Schwinger-Dyson equations in our discretized world will be a set of first-order PDEs. Nominating our Fourier transform variable to be ϕ_n, we write the solution as

$$Z[J] \equiv \int [d\phi]\exp\left(iS[\phi] + iJ_n\phi_n\right) \tag{309}$$

The Fourier-transformed Schwinger-Dyson equations then become a differential equation for this creature S:-

$$\frac{\partial S}{\partial\phi_m} = K_{mn}\phi_n - \frac{\partial V}{\partial\phi_n} \tag{310}$$

This equation has an obvious solution:

$$S = \tfrac{1}{2}\phi_m K_{mn}\phi_n - V(\phi) + c \tag{311}$$

The additive constant cannot be determined from the Schwinger-Dyson equations as those equations are homogeneous in Z. Let us now return to the continuum limit. We have:

$$S = \int d^Dx\left[-\tfrac{1}{2}\phi\partial^2\phi - V(\phi)\right] = \int d^Dx\left[+\tfrac{1}{2}(\partial\phi)^2 - V(\phi)\right] \tag{312}$$

where we used integration by parts and absorbed a normalization constant.

Notice that this is the action for the scalar field theory whose physics we started out to study. In other words, our friend the familiar Feynman Path Integral with a source term

$$S_{\rm source} = \int d^Dx\,\phi(x)\cdot J(x) \tag{313}$$

literally is the object which is the generating functional for *all* Green's functions - in the [free or] interacting theory. This fact holds in canonical quantization so of course it must be true! This gives us another reason, albeit based in canonical quantization, to pay attention to the generating functional with source $Z[J]$. Either way, $Z[J]$ is a very powerful animal and we will be using it to make sense of loop diagrams for spins $0, \frac{1}{2}, 1$ in the path integral approach.

5.8 The S-matrix and the LSZ reduction formula

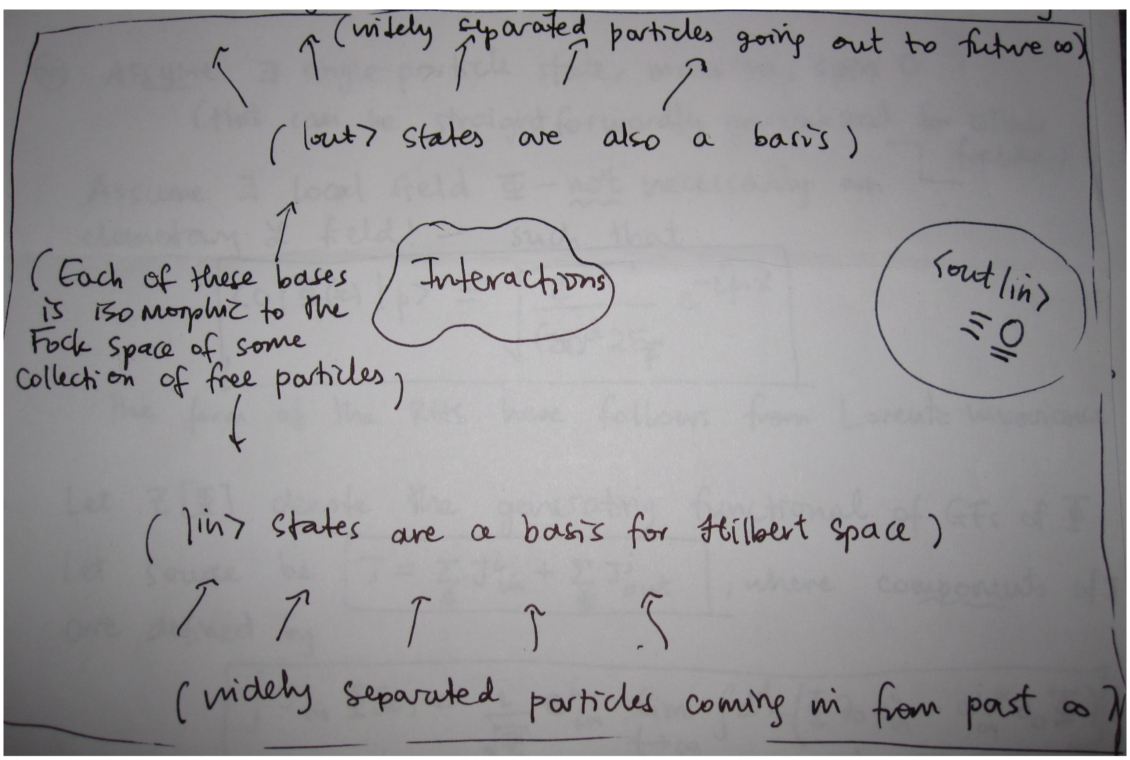
This subsection is based on part of the slim but powerful textbook by Thomas Banks.

Consider a situation in which past infinity and future infinity are places where we can widely separate incoming/outgoing particles. Define a basis of "in-states" $|\text{in}\rangle$ and a different basis of "out-states" $|\text{out}\rangle$, which each form a basis isomorphic to the Fock space of a collection of free particles. The overlap is $\langle\text{in}|\text{out}\rangle = 0$. The interior of spacetime is the place where the all-important interactions occur. The *S-matrix* is defined to be the animal which takes in-states to out-states.

One very important property of QFT is the *cluster property*, which says that connected parts of Green's functions fall to zero at large spacelike separations. For massive fields this falloff is exponential, making definition of in- and out-states straightforward. For massless fields it is power-law (think Coulomb's Law), and because of infrared (IR) divergences the business of defining an S-matrix can be tricky. In the following, we pick massive fields, in order to sidestep this technical subtlety. Further information may be found in the textbooks listed on the course web site.

The S-matrix is a unitary transformation on a subspace of Hilbert space, the "scattering states", which are orthogonal to all bound states of the particles. The entire Hilbert space is thus spanned by scattering states, but only if we treat bound states as separate particles. We will make this assumption in the following as well.

Here is a cartoon picture of the setup for the S-matrix.



In the interacting QFT it is very difficult to compute the physics and know the energy eigenstates of the Hamiltonian. The technology of Green's Functions allows us to finesse this problem somewhat. But our most helpful tool here will actually be the Lehmann-Symanzik-Zimmerman (LSZ) reduction formula, which provides us the connection between Green's functions and Feynman graph technology and the physical S-matrix.

Assume that there exists a single-particle state, of mass m and spin zero. (Note: this story can be straightforwardly generalized for other types of fields than massive scalars.) Assume further that there exists a local field Φ, which is not necessarily an elementary Lagrangian field, such that

$$\langle 0|\Phi(x)|p\rangle = \sqrt{\frac{z}{(2\pi)^d 2E_p}} e^{ip\cdot x} \tag{314}$$

where E_p is the energy associated to the particle with momentum $\vec{p}$. As you should check, the form of the RHS here follows directly from Lorentz invariance. z is a normalization constant and should not be mistaken for the entire functional integral $Z[J]$.

Let $Z[\Phi]$ denote the generating functional of Green's functions of the field Φ. Let the source be

$$J = \sum_i J^i_{\text{in}} + \sum_j J^j_{\text{out}} \tag{315}$$

where components of J are defined by

$$\int J^i_{\text{in}}\Phi(x) = \frac{i}{\sqrt{z}}\epsilon^i_{\text{in}} \lim_{t\to-\infty} \int d^d x \left(\Phi\partial_0\phi^i_{\text{in}} - \phi^{i*}_{\text{in}}\partial_0\Phi\right) \tag{316}$$

Here ϵ^i_{in} is an infinitesimal expansion parameter, and ϕ^i_{in} is a normalizable positive-energy solution of the Klein-Gordon equation. For a free field, the RHS of the above equation would be just the creation operator for a one-particle wavefunction ϕ_{in}. Notice that

- For any local field $\Phi(x)$, the source creates a state localized around an infinitely spatially distant point in the limit $t\to-\infty$.
- For a matrix element $\langle\eta|\int J\Phi|\psi\rangle$ between states $|\psi\rangle$ and $|\eta\rangle$ with fixed momentum vectors p^μ, the only surviving terms in the limit $t\to-\infty$ arise from $(p^\mu_\psi - p^\mu_\eta)^2 = m^2$.
- If we pick all incoming and outgoing wavefunctions to be spatially localized around different asymptotic directions in the limit $t\to\pm\infty$, then this operator acts just like an "in" creation operator even in the *interacting* theory!

Similarly, the "out" part of the source J acts like a sum of annihilation operators. In that case, the outgoing wavefunctions are negative-energy solutions of the KG equation (this is a convention: we have chosen "all momenta incoming"), and we take the limit $t\to+\infty$ instead.

Now, from its definition, $Z[J]$ is just

$$Z[J] = \langle 0|T\exp(i\int J\cdot\phi)|0\rangle \tag{317}$$

In this expression for $Z[J]$, all annihilation operators sit to the left of all creation operators. So it is at least possible that $Z[J]$ with this source could be the generating functional of the

S-matrix. More precisely, the coefficient of $\epsilon_{\text{in}}^1 \cdots \epsilon_{\text{in}}^m \epsilon_{\text{out}}^1 \cdots \epsilon_{\text{out}}^n$ in the expansion of $Z[J]$ with this source is the S-matrix element

$$\langle \text{out}; f_1 \cdots f_n | g_1 \cdots g_m; \text{in} \rangle \tag{318}$$

Other terms in this expansion of the composite source are of no particular use; they just correspond to creation/annihilation of *multiple* particles in the same scattering state. In these expressions, the single-particle states f_i, g_j are defined by

$$\begin{aligned} \phi_{\text{in}}^i &= \int \frac{d^d p}{\sqrt{2E_p (2\pi)^d}} e^{-iE_p t - i\vec{p}\cdot\vec{x}} f_i(p) \\ \phi_{\text{out}}^j &= \int \frac{d^d p}{\sqrt{2E_p (2\pi)^d}} e^{iE_p t - i\vec{p}\cdot\vec{x}} g_j(p) \end{aligned} \tag{319}$$

Consider our "in" source term:

$$\int dt\, i\epsilon_{\text{in}}^i \partial_0 \left\{ \int d^d x \left(\Phi \partial_0 \phi_{\text{in}}^i - \phi_{\text{in}}^{i*} \partial_0 \Phi \right) \right\} \tag{320}$$

The boundary term at $t \to -\infty$ is what we needed, but not so for the case for the boundary term at $t \to +\infty$! The problem is that it contains a creation operator for a state localized in the *future*. The saviour of our sanity here is that all final states are orthogonal to all initial states, by the orthogonality of the "in" and "out" bases. So the wayward future creation operator just ends up getting killed off and having no practical effect. (For more on the case of partially forward scattering, see Banks and references.)

Our next step is to massage this source a bit, and for this we need three things. Firstly, we use an [outer] partial time derivative on the Klein-Gordon scalar product. Secondly, we make use of the Wronskian form. Thirdly, we use the fact that the ϕ_{in}^i obey the Klein-Gordon equation. (Note: Φ, which is not necessarily a fundamental Lagrangian field, typically does *not* obey the KG equation.) Making use of these three steps yields, as the interested reader should verify explicitly,

$$i \int d^D x \left\{ (\nabla^2 - m^2) \phi_{\text{in}}^i \Phi - \partial_0^2 \Phi \phi_{\text{in}}^i \right\} = i \int d^D x \phi_{\text{in}}^i \left(\partial^2 + m^2 \right) \Phi \tag{321}$$

The story for each of the other "in"-sources is morally identical.

The story for the out-sources the computation is very similar, with only straightforward changes like "in" $\longrightarrow$ "out", $t \to -\infty \longrightarrow t \to +\infty$.

The net result is the LSZ Reduction Formula:

$$\begin{aligned} &\langle \text{out}; f_1 \cdots f_n | g_1 \cdots g_m; \text{in} \rangle \\ &= \left(\frac{i}{\sqrt{z}} \right)^{m+n} \int \prod_k d^D x_k \phi_{\text{in}}^k (x_k) \prod_j d^D y_j (\phi_{\text{out}}^j)^* (y_j) \times \\ &\quad \times \left(\partial_{x_k}^2 + m^2 \right) \left(\partial_{y_j}^2 + m^2 \right) \times \langle 0 | T \left\{ \Phi(x_1) \cdots \Phi(x_n) \Phi(y_1) \cdots \Phi(y_m) \right\} | 0 \rangle \end{aligned} \tag{322}$$

Warning: you will often see $\phi_{\text{in}}, \phi_{\text{out}}$ replaced by plane waves. This creates some trivial infinities in cross-sections. These are easily dealt with via regularization such as spacetime discretization.

Notice that the sole requirement on $\Phi(x)$ is that it have finite amplitude to create single-particle states. This can be verified by looking for a pole in the two-point function.

A very important property of the LSZ reduction formula is that it is symmetric between in- and out-states – *except* that in-states have positive energy while out-states have negative energy. This remarkable fact suggests a symmetry of the physics known as **crossing symmetry**. The part of this which is technically difficult to prove is analyticity.

It is handy to have the momentum-space version of the LSZ Reduction Formula as well:

$$\begin{aligned}&\langle \text{out}; q_1 \cdots q_n | p_1 \cdots p_m ; \text{in} \rangle \\ &= \left(\frac{i}{\sqrt{z}} \right)^{m+n} \prod_i \frac{1}{\sqrt{2E_p (2\pi)^d}} \prod_j \frac{1}{\sqrt{2E_p (2\pi)^d}} \times \\ &\quad \times \left(p_i^2 - m^2 \right) \left(q_j^2 - m^2 \right) G_{n+m}(-q_1, \cdots, -q_m; p_1, \cdots, p_n) \end{aligned} \tag{323}$$

where G_{n+m} is the Fourier transform of the time-ordered product of $n + m$ Φ fields.

Nota Bene: For outgoing states, q_i is the physical positive-energy relativistic momentum *but* the Fourier Transform is evaluated at *negative* energy outgoing momenta $\{-q_j\}$

As mentioned briefly above, translation invariance is something we need to watch in the LSZ context. In particular, we will get a momentum-conserving $\delta^D(\sum_i p_i - \sum_j q_j)$. Squaring this on the way to finding the probability amplitude and thereby the scattering cross-section will result in an overall $\delta^D(0)$ term, which should be interpreted as the spacetime volume. This result eventuated because, in a translation-invariant system, we should really be after the probability *per unit volume*. With wavepackets rather than plane waves, this problem vanishes from the radar screen.

One of the most important things to note about the LSZ Reduction Formula is that all momenta are on the mass shell. Therefore, the formula involves a product of a bunch of zeroes! This is not a contradiction; the S-matrix element does physically vanish unless a Green's function has a single pole at the mass shell on each external leg. As we have seen for free scalar field theory and will see in some detail for more interesting cases shortly, this happens readily in interesting QFTs.

Earlier this week, we learned that the full Green's functions evaluate as trees whose vertices are the 1PI functions and whose external *and* internal lines are full propagators $W_2(p)$. As the next subsection shows, something called the Källen-Lehmann representation for propagators shows that each Green's function W_2 does indeed have a simple pole, yielding precisely the multiple-pole structure that is needed in order for the LSZ Reduction Formula to predict a finite S-matrix.

5.9 Källen-Lehmann spectral representation for interacting QFTs

Recall the completeness relation for one-particle states:

$$\mathbb{1}_{1-\text{particle}} = \int \frac{d^d p}{(2\pi)^d} \frac{1}{2E_p} |\vec{p}\rangle\langle\vec{p}| \tag{324}$$

Let $|\lambda_0\rangle$ be the eigenstate of the Hamiltonian with zero momentum:

$$\vec{p}|\lambda_0\rangle = \vec{0}|\lambda_0\rangle \tag{325}$$

Define

$$E_p(\lambda) \equiv \sqrt{|\vec{p}|^2 + m_\lambda^2} \tag{326}$$

where m_λ is the "mass" of the state $|\lambda_{\vec{p}}\rangle$, i.e. $E(|\lambda_0\rangle)$. Then the *full Hilbert space* completeness relation is

$$\mathbb{1} = |\Omega\rangle\langle\Omega| + \sum_\lambda \int \frac{d^dp}{(2\pi)^d} \frac{1}{2E_{\vec{p}}(\lambda)} |\lambda_{\vec{p}}\rangle\langle\lambda_{\vec{p}}| \tag{327}$$

We will now make use of this little beastie by inserting it in between operators at spacetime points x, y in the two-point function, in order to learn about the energy distribution of contributions to the propagator. We first do the analysis for the case $x^0 > y^0$; the analysis is very similar for $x^0 < y^0$.

First, notice that

$$\langle\Omega|\phi(x)|\Omega\rangle\langle\Omega|\phi(y)|\Omega\rangle = 0 \tag{328}$$

by symmetry (or by Lorentz invariance for spin $s > 0$). Therefore,

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle = \sum_\lambda \int \frac{d^dp}{(2\pi)^d} \frac{1}{2E_{\vec{p}}(\lambda)} \langle\Omega|\phi(x)\lambda_{\vec{p}}\rangle\langle\lambda_{\vec{p}}|\phi(y)|\Omega\rangle \tag{329}$$

Note that nothing here requires the ϕ field to be elementary. This makes the Källen-Lehmann representation a very useful one in a wide variety of QFT contexts, for example with composite operators in QCD.

In order to massage our expression further, we need to recall how translation generators act. For a translation by x,

$$\begin{aligned}
\langle\Omega|\phi(x)|\lambda_{\vec{p}}\rangle &= \langle\Omega|e^{i\hat{p}\cdot x}\phi(0)e^{-i\hat{p}\cdot x}|\lambda_{\vec{p}}\rangle \\
&= \langle\Omega|\phi(0)|\lambda_{\vec{p}}\rangle e^{-ip\cdot x}\big|_{p^0=E_{\vec{p}}} \\
&= \langle\Omega|\phi(0)|\lambda_{\vec{0}}\rangle e^{-ip\cdot x}\big|_{p^0=E_{\vec{p}}}
\end{aligned} \tag{330}$$

by Lorentz invariance of the vacuum $|\Omega\rangle$ and of $\phi(0)$. Therefore,

$$\langle\Omega|\phi(x)\phi(y)|\Omega\rangle|_{x^0>y^0} = \sum_\lambda \int \frac{d^Dp}{(2\pi)^D} \frac{ie^{-ip\cdot(x-y)}}{(k^2 - m_\lambda^2 + i\epsilon)} |\langle\Omega|\phi(0)|\lambda_{\vec{p}}\rangle|^2 \tag{331}$$

Similarly for the other case $x^0 < y^0$. Putting them together we can summarize the information via

$$\langle\Omega|T\{\phi(x)\phi(y)\}|\Omega\rangle = \int_0^\infty \frac{d\mu^2}{2\pi} \rho(\mu^2) \mathscr{D}_F(x-y;\mu^2) \tag{332}$$

where the *spectral density* function is given by

$$\rho(\mu^2) = \sum_\lambda (2\pi)\delta(\mu^2 - m_\lambda^2) \left|\langle\Omega|\phi(0)|\lambda_{\vec{0}}\rangle\right|^2 \tag{333}$$

Peskin and Schroeder have a very nice pictorial representation of what $\rho(\mu^2))$ looks like in a typical quantum field theory.

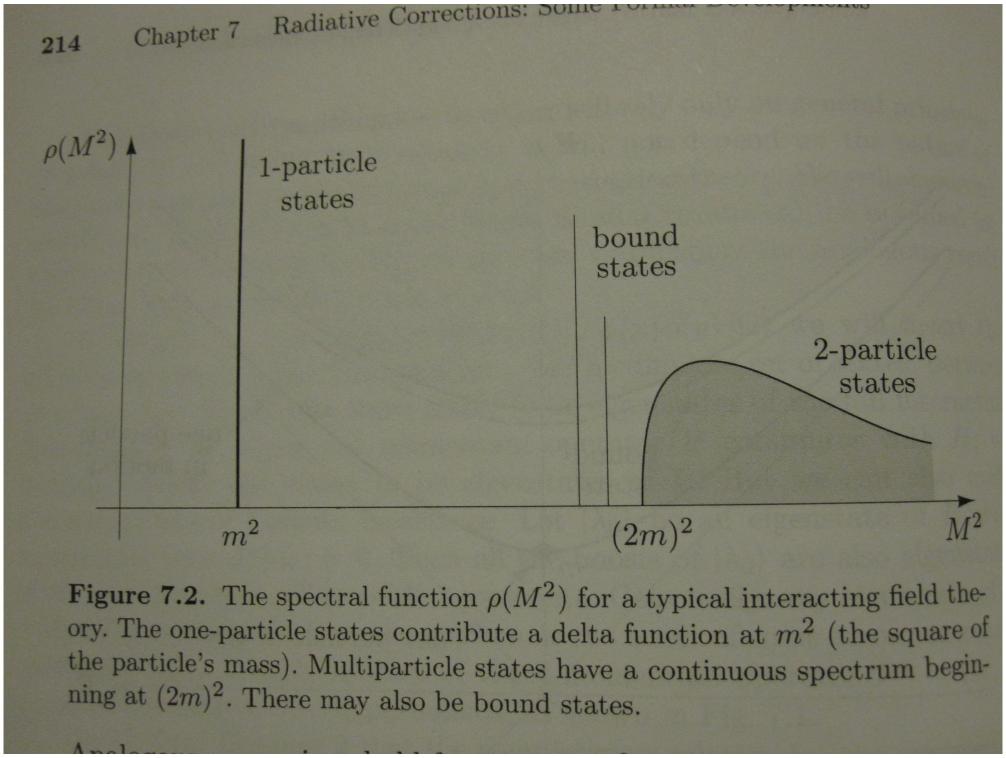


Figure 7.2. The spectral function $\rho(M^2)$ for a typical interacting field theory. The one-particle states contribute a delta function at m^2 (the square of the particle's mass). Multiparticle states have a continuous spectrum beginning at $(2m)^2$. There may also be bound states.

The main features are:

- There is a delta-function peak at $\mu^2 = m^2$, representing one-particle states.
- There are smaller peaks at $m^2 < \mu^2 < (2m)^2$, representing boundstates.
- At $(2m)^2$ the spectral density rises (with discontinuous derivative), peaks a bit later along the μ^2 axis, and then falls off again at larger μ^2.

Note that while the one-particle state gives an isolated delta function contribution, the two-particle spectrum is continuous. As you can see for yourself, this is a simple consequence of energy-momentum conservation.

The analytic structure can be easier to visualize in momentum space.

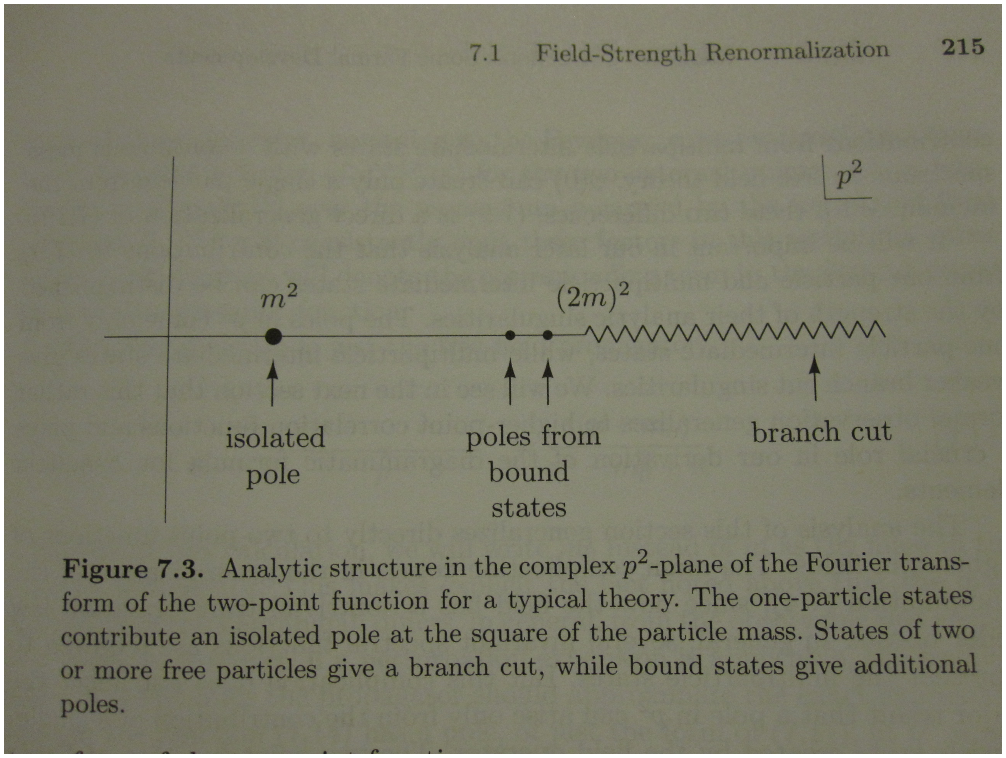


Figure 7.3. Analytic structure in the complex p^2-plane of the Fourier transform of the two-point function for a typical theory. The one-particle states contribute an isolated pole at the square of the particle mass. States of two or more free particles give a branch cut, while bound states give additional poles.

Since the relevant variable is momentum-squared here, we find that in the complex p^2 plane

- There is an isolated single-particle pole at $p^2 = m^2$, which is on the real axis.
- Poles arise between $m^2 < p^2 < (2m)^2$, corresponding to boundstates.
- At $p^2 = (2m)^2$ there is a *branch cut* singularity. As you will recall from your complex analysis class, this is a weaker type of singularity than a pole. Physically, the threshold at $p^2 = (2m)^2$ corresponds to creation of particle-antiparticle pairs. Below that, the pairs would only be virtual and so there is no corresponding branch cut structure for $p^2 < (2m)^2$.

The one-particle pole is isolated, at m^2.

The main physics lesson that can be drawn from this is that QFT is unequivocally and inherently a multi-particle animal. Life in relativistic quantum field theories is definitely not all about boring one-particle states!

None of the analysis here requires that the ϕ field be elementary, or that perturbation theory be valid. This is what makes it so powerful and useful.

6 Functional quantization for spin half

6.1 FPI quantization for fermions: Grassmann variables

Our development of functional quantization thus far has subtly depended upon thinking of the fields as mathematical objects with bosonic statistics. The spin-statistics theorem says that free spin-half fields, on the other hand, must obey the Pauli principle. So how are we to incorporate fermionic behaviour into the formalism of path integrals?

It turns out to be surprisingly easy to solve this apparent conundrum. We introduce a new concept called **Grassmann** variables, which *anti*commute. For any two such anticommuting fellows θ, η we have

$$\{\theta, \eta\} = 0 \tag{334}$$

In other words, $\theta\eta = -\eta\theta$. Notice that a fermion bilinear is again a boson. Since the above equation holds for any θ, η, it holds in particular when $\theta = \eta$, i.e.

$$\theta^2 \equiv 0 \tag{335}$$

This mathematizes the Pauli principle. It also makes Taylor expansions splendidly easy, because each such Taylor series terminates after the linear piece:

$$f(\theta) = A + B\theta \tag{336}$$

where A, B are constants.

An immediate question comes to mind after defining Grassmann variables: how do we define integration? Differentiation? The most physically useful definition for integration, also motivated by mathematics, has proven to be

$$\begin{aligned} \int d\theta &= 0 \\ \int d\theta\theta &= 1 \end{aligned} \tag{337}$$

In cases where there are multiple θ^α, such as would occur in models with extended supersymmetry, we would have

$$\int d\theta^\alpha \theta_\beta = \delta^\alpha{}_\beta \tag{338}$$

How about the *ordering* of integration? We do have to be extremely careful about minus signs for Grassmann variables, because they anticommute. Define

$$\int d\theta \int d\eta \cdot \eta \cdot \theta = +1 \tag{339}$$

This definition amounts to a choice of sign convention, but it matters physically that we make this same choice consistently.

How about complex Grassmann variables?

$$(\theta\eta)^* = \eta^*\theta^* \tag{340}$$

This behaviour is easy to remember because it smells a lot like matrices. Also,

$$\int d\theta^* d\theta\, \theta\, \theta^* = +1 \tag{341}$$

How about derivatives? How do we define the *ordering* of a Grassmann derivative, consistent with the anticommutation property of Grassmann variables? We will pick the same convention as Peskin & Schroeder: the one in which derivatives are defined as **left derivatives**:

$$\frac{d}{d\eta}(\theta\eta) = -\frac{d}{d\eta}(\eta\theta) = -\theta \tag{342}$$

Note that $\partial/\partial\eta$ is a Grassmann object itself, so must obey the usual anticommutation property in addition to behaving like a derivative.

One of the most fun parts of using Grassmann fields for modelling relativistic fermion fields is realizing how dead-simple Gaussians become. Because Taylor expansions truncate so early owing to the anticommutation property of Grassmann fields, we have

$$\exp\left(-\theta^* b\theta\right) = 1 - \theta^* b\theta + 0 \tag{343}$$

Therefore,

$$\int d\theta^* d\theta e^{-\theta^* b\theta} = \int d\theta^* d\theta \left(1 - \theta^* b\theta\right) = b \tag{344}$$

by the rules of Grassmann integration and the anticommuting property.

In doing Feynman path integrals to find correlation functions of physical fields, we needed to insert fields into the integrand. So let us consider for instance

$$\begin{aligned} & \int d\theta^* d\theta\, \theta\, \theta^* e^{-\theta^* b\theta} \\ = & \int d\theta^* d\theta\, \theta\, \theta^* \left(1 - \theta^* b\theta\right) \\ = & \int d\theta^* d\theta\, \theta\, \theta^* = 1 \end{aligned} \tag{345}$$

Compare this to the result we had obtained just above. Looking carefully, we notice that the Gaussian integral with $\theta\theta^*$ in the integrand brings down an extra factor of $(1/b)$ compared to the case without. This echoes closely what happens for scalar fields, as you can check by reminding yourself of how to perform bosonic Gaussian integrals.

There is a spin-half fermion analogue of the functional determinant formula we obtained earlier for spin-zero bosons, so let us derive it. Consider the following quantity for an invertible matrix B_{ij}:

$$\left(\prod_i \int d\theta_i^* d\theta_i\right) \exp\left(-\theta_i^* B_{ij}\theta_j\right) \tag{346}$$

In the diagonal basis, B_{ij} has eigenvalues $\{b_i\}$. Therefore, the value of the above quantity is

$$\prod_i \int d\theta_i^* d\theta_i \exp\left(-\sum_i \theta_i^* b_i \theta_i\right) = \prod_i b_i = \det(B) \tag{347}$$

I have written out the Einstein summation convention indices explicitly here, as the meaning of the repeated indices could otherwise be read as ambiguous.

We can rewrite this formula as

$$\prod_i \int d\theta_i^* d\theta_i e^{-\theta^* B\theta} = [\det(B)]^{+1} \tag{348}$$

Notice that something remarkable has happened here: we have obtained a *very* similar expression to what we found for the functional determinant involving scalar fields! The only difference is that the determinant appears to the +1 power rather than $-\frac{1}{2}$.

As you should explicitly check, it is straightforward to show that with multiple θ_j the analogous expression is

$$\left(\prod_i \int d\theta_i^* d\theta_i\right) \theta_k \theta_\ell^* \exp\left(-\theta_i^* B_{ij} \theta_j\right) = [\det(B)]^{+1} \left(B^{-1}\right)_{k\ell} \tag{349}$$

The main idea that goes into showing this is just the paragraph at the end of the previous subsection.

Another interesting formula may be obtained by considering an $N \times N$ antisymmetric matrix A and evaluating

$$\int d^n\eta \, e^{\frac{1}{2}A_{ij}\eta_i\eta_j} = \epsilon_{a_1\cdots a_N} A_{a_1a_2} \cdots A_{a_{N-1N}} \equiv \mathrm{Pf}[A] \tag{350}$$

where Pf$[A]$ is known as the *Pfaffian* of the antisymmetric matrix A. It is defined to be zero if N is odd but, as we see here, is definitely nonzero for even N.

6.2 Generating functional with sources for Dirac fields

From earlier in our physics training, we are familiar with Dirac fields $\psi(x)$ satisfying

$$(i\partial\!\!\!/ - m)\,\psi = 0 \tag{351}$$

which follows from the Dirac action

$$S = \int d^Dx \left[\bar\psi\,(i\partial\!\!\!/ - m)\,\psi\right] = \int d^Dx \left[\frac{i}{2}\bar\psi \overleftrightarrow{\partial\!\!\!/}\psi - m\bar\psi\psi\right] \tag{352}$$

where in the second expression we used an integration by parts to symmetrize the action in $\bar\psi, \psi$.

In order to write down the fermionic Feynman path integral, we will need to make $\psi(x)$ a Grassmann *field*:

$$\psi(x) := \sum_i \psi_i f_i(x) \tag{353}$$

where ψ_i are Grassmann number coefficients and $\{f_i(x)\}$ are a set of spinor-valued basis functions of spacetime coordinates.

This expansion is fundamentally different than what you write down in canonical quantization, where the field is expanded in terms of Fourier modes with operator coefficients obeying canonical anticommutation relations. Here, we have Grassmann number coefficients, *not* operator coefficients.

The definition of correlation functions for spin-$\frac{1}{2}$ fermions goes over in pretty direct analogy to what you would expect based on experience with FPIs for scalar fields. There are in particular no gauge freedom subtleties here, so we do not yet have to touch the concept of Fadeev-Popov ghosts.

Let us consider the two-point function

$$\langle 0|T\left\{\psi(x_1)\bar{\psi}(x_2)\right\}|0\rangle = \frac{\int \mathscr{D}\bar{\psi}\mathscr{D}\psi\,\psi(x_1)\bar{\psi}(x_2)\,\exp\left[i\int d^Dx\,\bar{\psi}\,(i\partial\!\!\!/ - m)\,\psi\right]}{\int \mathscr{D}\bar{\psi}\mathscr{D}\psi\exp\left[i\int d^Dx\,\bar{\psi}\,(i\partial\!\!\!/ - m)\,\psi\right]} \tag{354}$$

This can be evaluated using our formulæ above, including (349). In Fourier space, the result is:

$$\begin{aligned}\langle 0|T\left\{\psi(x_1)\bar{\psi}(x_2)\right\}|0\rangle &= S_F(x_1-x_2)\\ &= \int \frac{d^Dk}{(2\pi)^D}\frac{i\,e^{-ik\cdot(x_1-x_2)}}{(k\!\!\!/ - m + i\epsilon)}\end{aligned} \tag{355}$$

Suppose we are interested in doing rather more than just simple free field theory for our fermionic spin-half fields. As we found for scalar fields, by far the simplest way to extract the correlation functions is to use the method of the *functional source*. We therefore need to ape the source concept for fermions.

Since any animal appearing in the action (a scalar under Poincaré group transformations) had better be bosonic in nature, we define a new fermionic generating functional with *fermionic* sources for $\psi, \bar{\psi}$ denoted $\bar{\eta}, \eta$:

$$\begin{aligned}Z[\bar{\eta},\eta] &= \int \mathscr{D}\bar{\psi}\psi\,\exp\left[i\int d^Dx\left\{\bar{\psi}\,(i\partial\!\!\!/ - m)\,\psi + \bar{\eta}\psi + \bar{\psi}\eta\right\}\right]\\ \text{i.e. } Z[\bar{\eta},\eta] &= \int \mathscr{D}\bar{\psi}\mathscr{D}\psi e^{iS[\bar{\psi},\psi]}e^{i\int d^Dx\left(\bar{\eta}\psi+\bar{\psi}\eta\right)}\end{aligned} \tag{356}$$

What is the physical reason why we need sources for both ψ and $\bar{\psi}$? Does that not constitute overcounting? No, it does not. In four dimensions, Dirac fermions have four off-shell and two on-shell components. (In other dimensions, we get a different power of two, $2^{[D/2]}$, halved on-shell.) Since we plan to be handling full off-shell physics at quantum loop level in perturbation theory, we do not want to make any silly assumptions about fields appearing in loops being on-shell. That is why we carry around the full $\eta, \bar{\eta}$ baggage: it is physically essential.

As for scalar fields, note that this generating functional $Z[\bar{\eta},\eta]$ is linear in the sources. The mathematical bonus of this physically motivated definition is that we can easily go about evaluating it just by completing the square! Check for yourself that the answer is:

$$Z[\bar{\eta},\eta] = Z_0\exp\left(-\int\int d^Dx\,d^Dy\,\bar{\eta}(x)S_F(x-y)\eta(y)\right) \tag{357}$$

where Z_0 is defined to be the generating functional evaluated at $\bar{\eta}=0$ and $\eta=0$.

Using information from earlier in this section concerning Grassmann derivatives, we have, mimicking the scalar case,

$$\langle 0|T\left\{\psi(x_1)\bar{\psi}(x_2)\right\}|0\rangle = \left\{\frac{1}{Z_0}\left(-i\frac{\delta}{\delta\bar{\eta}(x_1)}\right)\left(+i\frac{\delta}{\delta\eta(x_2)}\right)Z[\bar{\eta},\eta]\right\}\Bigg|_{\bar{\eta}=0\,,\eta=0} \tag{358}$$

Note the $-i$ in the factor $-i\delta/\delta\bar{\eta}(x_1)$ which pulls down a $\psi(x_1)$; that is necessary to keep our definitions of Grassmann integrals and derivatives consistent. Likewise with the $+i$ in the factor $+i\delta/\delta\eta(x_2)$ which pulls down a $\bar{\psi}(x_2)$ from the exponential. Higher correlation functions would of course also be obtained by iterating this concept.

6.3 Weyl and Majorana fields

Weyl fermions are chiral. They can be defined in any even dimensional spacetime, where it is possible to define γ_{D+1} in analogy to how we defined γ_5 in $D=4$. (In odd dimensions, the product of all γ matrices ends up being proportional to the identity, by Schur's Lemma.) We will now restrict to $D=4$ to save notational baggage. Nota bene: interested readers should consult e.g. the relevant appendix in the Polchinski string theory texts for a great rendition of how to analyze fermion representations in arbitrary spacetime dimension D. The oscillator representation discussed there, built out of pairs of γ matrices and their anticommutation relations, provides a nice physical way to understand why the general fermion representations have dimensionality $2^{[D/2]}$. Putting the restriction on the fermion that it be chiral reduces the dimensionality to the Weyl one: $2^{[D/2]-1}$.

Consider a left-handed Weyl field; this is a representation of Lorentz which transforms as $(\frac{1}{2},0)$ under $SU(2)_L\times SU(2)_R$. It is notationally traditional to denote this field as ν_a. Right-handed Weyl fields transform as $(0,\frac{1}{2})$ and have indices that are traditionally dotted: $\nu_{\dot{a}}$. As you should check, the Levi-Civita symbol provides a Lorentz-invariant product: $\nu_a\nu_b\epsilon^{ab}$. Because of this, we can raise and lower indices using ϵ:

$$\begin{aligned} \epsilon^{ab}\nu_b &= \nu^a \\ \epsilon_{ab}\nu^b &= \nu_a \\ \epsilon^{ab}\epsilon_{bc} &= \delta^a_{\ c} \end{aligned} \tag{359}$$

This is specific to $SU(2)$.

Suppose we want to tensor two Weyl representations together in an attempt to build a Lorentz-invariant action involving derivatives. In that case, we would need a Clebsch-Gordan coefficient to connect the two Weyl representations to the vector $(\frac{1}{2},\frac{1}{2})$ representation. The C-G roles are played by the following vectors which are defined in terms of Pauli σ^i matrices as

$$\begin{aligned} \left\{\sigma^{\mu}_{a\dot{b}}\right\} &= \left\{\mathbb{1},+\vec{\sigma}\right\}_{a\dot{b}} \\ \left\{\bar{\sigma}^{\mu}_{\dot{b}a}\right\} &= \left\{\mathbb{1},-\vec{\sigma}\right\}_{\dot{b}a} \end{aligned} \tag{360}$$

This permits writing down the Weyl lagrangian:

$$\mathscr{L}_W = \frac{1}{2}i(\nu^*)^{\dot{b}}(\bar{\sigma}^{\mu})_{\dot{b}a}\partial_{\mu}\nu^a + \text{h.c.} \tag{361}$$

The story for the right-handed Weyl fields is identical except with right-handed dotted spinors traded for left-handed undotted spinors and σ^μ traded for $\bar{\sigma}^\mu$.

Note that the Weyl lagrangian is invariant under a phase shift of ν. The corresponding conserved quantity is, for free particles, helicity.

Weyl fields are massless. We discussed why when we talked about the Pauli-Lubański vector W^μ, quadratic Casimirs of Lorentz $C_{1,2}$, the Little Group of p^μ and helicity h. Here, $h = \frac{1}{2}$.

Since the Dirac field is the direct sum $(0, \frac{1}{2}) \oplus (\frac{1}{2}, 0)$, it makes sense to define

$$(\gamma^\mu) \equiv \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \tag{362}$$

Majorana fields are real. The ability to define Majorana spinors is, however, not present in all spacetime dimensions. The mathematics behind how this works is known as the Clifford algebra. All physically important properties in signature (s,t) spacetime in $D = s + t$ dimensions turn out to follow from the structure of $(s-t)$ mod 8. A mathematically oriented discussion of this can be found in
`http://www.maths.ed.ac.uk/~jmf/Teaching/Lectures/Majorana.pdf`. I also liked a little textbook called "The Spinorial Chessboard" the summer before grad school.

There are different kinds of Majorana conditions. For definiteness, let us work in one timelike and d spacelike dimensions. Then the Majorana condition requires $(s - t)$ mod 8 $= 6, 7, 0$, while the symplectic Majorana condition requires $(s-t)$ mod 8 $= 2, 3, 4$. Majorana-Weyl spinors occur when both Weyl and Majorana spinors are defined, i.e. in $D = s + t$ such that $(s - t)$ mod 8 $= 0$. This includes the cases $D = 1 + 1$ and $D = 1 + 9$ relevant to worldsheet and spacetime formulations of string theory. Similarly, symplectic Majorana-Weyl spinors occur when both Weyl and Symplectic Majorana spinors are defined, i.e. when $(s - t)$ mod 8 $= 4$.

Table B.1. *Dimensions in which various conditions are allowed for $SO(d-1,1)$ spinors. A dash indicates that the condition cannot be imposed. For the Weyl representation, it is indicated whether these are conjugate to themselves or to each other (complex). The final column lists the smallest representation in each dimension, counting the number of real components. Except for the final column the properties depend only on d mod 8.*

d	Majorana	Weyl	Majorana–Weyl	min. rep.
2	yes	self	yes	1
3	yes	-	-	2
4	yes	complex	-	4
5	-	-	-	8
6	-	self	-	8
7	-	-	-	16
8	yes	complex	-	16
9	yes	-	-	16
10=2+8	yes	self	yes	16
11=3+8	yes	-	-	32
12=4+8	yes	complex	-	64

7 Functional Quantization for Spin One

Only for massless spin one gauge fields is there a gauge invariance of the free action. The massless case is therefore our focus, as ensuring the FPI is defined in the face of gauge invariance is more technical than doing it for the massive case.

7.1 U(1) QED and gauge invariance

Our basic quantum field for the spin-one massless gauge field is $A_\mu(x)$. The field strength $F_{\mu\nu}$ is *not* a fundamental Lagrangian field; it is A_μ alone which has that status. The field strength, a derived quantity, is defined for $U(1)$ by

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \tag{363}$$

which is gauge-invariant. The action for A_μ is

$$S_{U(1)} = \int d^D x \left(-\tfrac{1}{4} F^{\mu\nu} F_{\mu\nu}\right) \tag{364}$$

We can rearrange this action via an integration by parts, to make it easier to pick off the Feynman propagator for the photon by eye:

$$\begin{aligned} S_{U(1)} &= \int d^D x \left(-\tfrac{1}{4} F^{\mu\nu} F_{\mu\nu}\right) \\ &= \int d^D x \left[+\tfrac{1}{2} A_\mu \left\{\eta^{\mu\nu}\partial^2 - \partial^\mu \partial^\nu\right\} A_\nu\right] \end{aligned} \tag{365}$$

Since this is a teensy bit cumbersome in position space, we can rewrite it in momentum space:

$$S_{U(1)} = \int \frac{d^D k}{(2\pi)^D} \frac{1}{2} \tilde{A}_\mu(k) \left[-\eta^{\mu\nu} k^2 + k^\mu k^\nu\right] \tilde{A}_\nu(k) \tag{366}$$

Is this all we need to worry about? It turns out that the answer to this is an emphatic *no*. The reason for our need to be careful is (da-daaa!): gauge invariance.

Recall from your previous experience that the EM field has a special property: $S[A_\mu]$ is invariant under gauge transformations leaving the field strength invariant:

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \alpha(x) \tag{367}$$

where $\alpha(x)$ is an arbitrary scalar function. If we had matter field(s) Ψ of charge q coupled to our $U(1)$ gauge field, then they would be rotated under $U(1)$ gauge transformations as $\Psi \to e^{iq\alpha(x)}\Psi$.

The physical consequences of this gauge invariance are (at least!) twofold: (a) Green's functions become ill-defined, and (b) the FPI becomes ill-defined, because simply integrating over all field configurations naïvely overcounts by an infinite amount! Let us now work to see how this comes about. We will then end up this week's discussion by showing how to fix this overcounting and the gauge invariance in one fell swoop known as the Fadeev-Popov Procedure.

Let us look back to the form of the action we wrote above, (366). Notice that if

$$\tilde{A}(k) = c\, k_\nu\, \alpha(k) \tag{368}$$

then something catastrophic occurs: the operator $[-\eta^{\mu\nu}k^2 + k^\mu k^\nu]$ annihilates it! Expressing this another way, if we write the Green's function equation as

$$\left[-\eta^{\mu\nu}k^2 + k^\mu k^\nu\right] G^1_{\nu\lambda}(k) = i\delta^\mu{}_\lambda \tag{369}$$

then it is clear that the Green's function as written has zero eigenvalues. The matrix $[-\eta^{\mu\nu}k^2 + k^\mu k^\nu]$ is singular. *Oops!*

7.2 Fadeev-Popov Procedure: Abelian case

Fadeev and Popov realized that it is possible to write down a *well-defined* Feynman Path Integral for gauge fields, by using a clever mathematical trick which effectively "damps" the out-of-control overcounting emanating from gauge invariance, giving rise to something physically reasonable.

To begin the F-P story, we introduce a new animal known as the **gauge fixing function** $G(A)$. (Note that this G has nothing to do with the Green's function; you will know this by its index structure.) An example would be

$$G(A) = \partial_\mu A^\mu \qquad \text{in Lorentz gauge} \tag{370}$$

If we then insist on imposing $G(A) \equiv 0$, this will restrict field configurations in the path integral over field space to take account of only those gauge fields obeying the gauge condition. This will be expressed mathematically shortly via a $\delta(G(A))$.

Consider the following reprentation of $\mathbb{1}$:

$$\mathbb{1} = \int \mathscr{D}\alpha(x)\delta(G(A))\det\left(\frac{\delta G(A^{(\alpha)}}{\delta\alpha}\right) \tag{371}$$

where

$$A^{(\alpha)}_\mu(x) = A_\mu(x) - \partial_\mu\alpha(x) \tag{372}$$

This formula is just the generalization of the familiar fact that $\delta(ax) = (1/a)\delta(x)$.

We now insert this unity factor into the gauge field FPI. In Lorenz gauge, it is

$$\mathbb{1} = \int \mathscr{D}\alpha(x)\delta(G_L(A))\det\left(-\partial^2\right) \tag{373}$$

In obtaining this expression, we used the fact that

$$\det\left(\frac{\delta}{\delta\alpha}\partial^\mu A^{(\alpha)}_\mu\right) = \det\left(-\partial^\mu\partial_\mu\alpha\right) \tag{374}$$

The fortunate accident for $U(1)$ is that this expression is independent of the gauge field!

Therefore, for our FPI we can write

$$\begin{aligned} Z_{U(1)} &= \int \mathscr{D}\alpha \int \mathscr{D}A e^{iS[A]} \delta(G(A)) \det\left(\frac{\delta G(A^{(\alpha)})}{\delta\alpha}\right) \\ &= \int \mathscr{D}\alpha \det\left(\frac{\delta G(A^{(\alpha)})}{\delta\alpha}\right) \int \mathscr{D}A e^{iS[A]} \delta(G(A)) \end{aligned} \tag{375}$$

Now let us look back to our gauge transformation (367). Note that it is, in *field* terms, just a shift of the gauge field $A_\mu(x)$. Accordingly,

$$\mathscr{D}A = \mathscr{D}A^{(\alpha)} \tag{376}$$

Also, of course,

$$S[A^{(\alpha)}] = S[A] \tag{377}$$

by gauge invariance. Therefore, in our FPI, we could just as well pick $A_\mu^{(\alpha)}$ as our variable of functional integration as A_μ itself! This enables us to write

$$\begin{aligned} Z_{U(1)} &= \int \mathscr{D}\alpha \det\left(\frac{\delta G(A^{(\alpha)})}{\delta\alpha}\right) \int \mathscr{D}A^{(\alpha)}\, e^{iS[A^{(\alpha)}]}\, \delta(G(A^{(\alpha)})) \\ &= \int \mathscr{D}\alpha \det\left(\frac{\delta G(A)}{\delta\alpha}\right) \int \mathscr{D}A\, e^{iS[A]}\, \delta(G) \end{aligned} \tag{378}$$

as A is just a dummy variable of integration in the FPI.

Now, our determinant in Lorenz gauge does not actually depend on α, so the factor $\int \mathscr{D}\alpha$ becomes a multiplicative infinite constant out front. This will cancel in all physical correlation functions. This decoupling is special to $U(1)$; we will find that for all other (non-Abelian) gauge theories the Fadeev-Popov ghosts do not decouple from the physical scattering amplitudes. It is actually critical for gauge invariance that they be involved; they do the job of "bookkeeping".

The next step is a very important one. Peskin & Schroeder gloss over this point as if it is obvious, which it is not. What they decide to do, in order to write down a well-defined FPI for $U(1)$ gauge theory, is to restrict to

$$G(A) = \partial_\mu A^\mu - \omega(x) \tag{379}$$

and insert a Gaussian convergence factor of

$$\int \mathscr{D}\omega \exp\left(-i \int d^D x\, \frac{1}{2\xi}\, \omega^2\right) \tag{380}$$

into the FPI. The factor of $-i$ out front here is required in $D = d + 1$, as can be seen by starting from a familiar Gaussian damping factor in Euclidean space and Wick rotating to Minkowski signature. We will return to giving a more complete understanding of the origin this mysterious damping factor when we discuss Fadeev-Popov ghosts properly – once we get to doing the case of non-Abelian gauge theories properly. For now, though, we will just take the P&S version as a definition of the FPI for $U(1)$ gauge theory:

$$Z = N(\xi) \int \mathscr{D}\omega \exp\left(-i \int d^D x\, \frac{1}{2\xi}\, \omega^2\right) \times \det\left(-\partial^2\right) \times$$

$$\int \mathscr{D}\alpha \times \int \mathscr{D}A\, e^{iS[A]}\, \delta\left(\partial_\mu A^\mu(x) - \omega(x)\right)$$
$$= \left\{N(\xi)\det\left(-\partial^2\right)\int \mathscr{D}\alpha\right\}\int \mathscr{D}A \exp\left(-i\int d^Dx\left[\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2\right]\right) \tag{381}$$

where in the last step we evaluated the $\int \mathscr{D}\omega$ integral using the δ-function[al]. The constant piece is physically unimportant as it cancels out of all physical correlation functions.

The upshot of this Fadeev-Popov convergence factor 'trickery' is that all correlation functions get computed using not the original Maxwell action but the gauge-fixed action:

$$S_{\text{tot}} = S_{\text{Maxwell}} + S_{\text{gf}} \tag{382}$$

where the total action includes a gauge-fixing term for $A_\mu(x)$. Interestingly, and satisfyingly, this statement also holds true for all S-matrix elements. In Lorenz gauge, the gauge-fixing term in the action is

$$S_{\text{gf,L}} = \int d^Dx\left[-\frac{1}{2\xi}\left(\partial_\mu A^\mu\right)^2\right] \tag{383}$$

7.3 Gauge-Fixed Photon Action at Tree Level

Our action for the photon, in Lorenz gauge, becomes

$$\begin{aligned} S_{\text{tot}} &= S[A] + S_{\text{gf}} \\ &= \int d^Dx\left\{+\frac{1}{2}A_\mu(x)\left[\eta^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu\right]A_\nu(x) - \frac{1}{2\xi}\left(\partial^\mu A_\mu\right)^2\right\} \\ &= \frac{1}{2}\int \frac{d^Dk}{(2\pi)^D}A_\mu(k)\left[-\eta^{\mu\nu}k^2 + k^\mu k^\nu - \frac{1}{\xi}k^\mu k^\nu\right]A_\nu(k) \end{aligned} \tag{384}$$

Therefore, the equation satisfied by the Lorenz gauge Feynman propagator is

$$\left[-k^2\eta_{\mu\nu} + k_\mu k_\nu\left(1 - \frac{1}{\xi}\right)\right]\tilde{\mathscr{D}}_F{}^{\nu\lambda}(k) = i\delta_\mu{}^\lambda \tag{385}$$

This has the benefit of not being a singular beastie. The solution is

$$\tilde{\mathscr{D}}_F{}^{\mu\nu}(k) = \frac{-i}{k^2 + i\epsilon}\left[\eta^{\mu\nu} - (1-\xi)\frac{k^\mu k^\nu}{k^2}\right] \tag{386}$$

You should go through the algebra to check this equation, which is (I believe) typo-free. The easiest way to show this is to (a) define the Feynman propagator for the $U(1)$ gauge field A_μ as

$$\tilde{\mathscr{D}}_F(k) = \frac{-i}{k^2 + i\epsilon}\left[\eta^{\mu\nu} - f(\xi)\frac{k^\mu k^\nu}{k^2}\right] \tag{387}$$

and then (b) attack $\tilde{\mathscr{D}}_F$ with the operator $(-k^2\eta_{\mu\nu} + k_\mu k_\nu(1 - 1/\xi))$, hoping for $+i\delta_\mu{}^\lambda$, and then (c) go through the algebra to show that $f(\xi) = 1 - \xi$.

7.4 Functional quantization for Yang-Mills fields: quick version

For the case of $U(1)$ of EM, we have seen last time the danger of trying to directly path-integrate over all possible gauge field configurations: the FPI blows up because we overcount, foolishly treating gauge-transformed fields differently to the originals. In order to render the path integral sensible and tame any wild oscillatory phases, we found it necessary to insert a convergence factor. This was natural, if a tad hand-waving! We write

$$Z = \int \mathscr{D}A_\mu e^{iS[A]} \sim \int \mathscr{D}\bar{A}_\mu e^{iS[A]} \int \mathscr{D}\Lambda \tag{388}$$

i.e. $\{A_\mu\}$ is the class of all gauge potentials reachable from $\bar{A}_\mu$ via a gauge transformation parametrized by $\Lambda(x)$.

Following Peskin and Schroeder, we insert a convergence factor to tame the FPI,

$$\int \mathscr{D}\Lambda \to \int \mathscr{D}G \exp\left(-\frac{i}{2\alpha}G^2\right) = \int \mathscr{D}\Lambda \det\left(\frac{\partial G}{\partial \Lambda}\right) \exp\left(-\frac{i}{2\alpha}G^2\right) \tag{389}$$

The resulting FPI is then

$$Z_{\rm YM} = \int \mathscr{D}A_\mu \int \mathscr{D}\Lambda \exp\left(i\int \mathscr{L}_{\rm YM} + J^\mu A_\mu - \frac{1}{2\alpha}G^2\right) \det\left(\frac{\partial G}{\partial \Lambda}\right) \tag{390}$$

Let us temporarily suppress the source term for the sake of notational clarity.

How are we to cope with the functional determinant? Let us write

$$M = \frac{\partial G}{\partial \Lambda} \tag{391}$$

as the matrix whose determinant we need for the FPI. Now look very hard at this expression. It should eventually ring a bell.

Earlier, we derived from Grassmann integration the rule that

$$\int \mathscr{D}\eta \mathscr{D}\bar{\eta} \exp\left(i\int \bar{\eta}M\eta\right) = [\det(M)]^{+1} \tag{392}$$

This is the key: it means that we can now **write the icky Jacobian as a functional determinant for Grassmann 'ghost' fields** $\eta, \bar{\eta}$**!**

Putting these facts together, we obtain the full Feynman path integral for Yang-Mills fields

$$Z_{\rm YM} = \mathscr{N} \cdot \int \mathscr{D}A_\mu \mathscr{D}\eta \mathscr{D}\bar{\eta} \exp\left[i\int \left(\mathscr{L}_{\rm YM} - \frac{1}{2\alpha}G^2 - \bar{\eta}M\eta\right)\right] \tag{393}$$

In this expression, G is the gauge-fixing function. The term

$$\mathscr{L}_{\rm gh} = -\bar{\eta}M\eta \tag{394}$$

is known as the ghost kinetic term, while the term

$$\mathscr{L}_{\rm gf} = -\frac{1}{2\alpha}G^2 = -\frac{1}{2\alpha}G^A G_A \tag{395}$$

is known as the gauge fixing term. Both are needed in order that physical scattering amplitudes are unitary at quantum loop level.

An extremely important feature of the Fadeev-Popov ghost fields is that **ghosts are *scalars***: they don't transform under Poincaré transformations. Since ghosts are scalars and are Grassmann fields, they do not obey the spin-statistics theorem. This is not a problem physically – **the Fadeev-Popov ghosts are bookkeeping devices which appear only in *internal* quantum loops, *never in external legs***.

7.5 Fadeev-Popov ghosts for Yang-Mills fields

The equation

$$G^A(A^B_\mu) = 0 \tag{396}$$

defines the gauge choice. For a compact symmetry group $\mathscr{G}$, it is possible to define a gauge-invariant measure known as the *Hurwitz measure*, obeying

$$dg = d(g'g) \tag{397}$$

where $g, g' \in \mathscr{G}$. Physically, only the *compact* groups give rise to a sensible gauge-invariant measure. We do not consider non-compact gauge groups at all in this course. Now recall that under U transformations

$$A_\mu \to A'_\mu = UA_\mu U^{-1} - \frac{i}{g}(\partial U)U^{-1} \tag{398}$$

where

$$U = \exp\left(-i\Delta\omega^A T_A\right) \tag{399}$$

Infinitesimally,

$$U \simeq 1 - i\Delta\omega^A T_A + \mathscr{O}(\Delta\omega)^2 \tag{400}$$

so that (dg) for $g \simeq 1$ infinitesimally close to the identity has the form

$$(dg) = \prod_A \Delta\omega^A \equiv d\vec{\omega} \tag{401}$$

$$A^A_\mu \to (A^A_\mu)' = A^A_\mu + \partial_\mu \Delta\omega^A + f_{BC}{}^A A^B_\mu \Delta\omega^C \tag{402}$$

Now consider the quantity $\Delta^{-1}[A]$ defined by

$$\Delta^{-1}[A] = \int \mathscr{D}\vec{\omega}\, \delta[G(A_\omega)] = \int \mathscr{D}g\, \delta[G[A_g]] \tag{403}$$

where

$$\begin{aligned} \mathscr{D}\vec{\omega} &= \prod_x d\omega(\vec{x}) \\ \mathscr{D}g &= \prod_x dg(x) \end{aligned} \tag{404}$$

Notice that Δ^{-1} is gauge invariant

$$\begin{aligned}\Delta^{-1}[A_g] &= \int \mathscr{D}g' \delta[G[A_{gg'}]] \\ &= \int \mathscr{D}(g'g)\delta[G[A_{g'g}]] \\ &= \int \mathscr{D}(g'')\delta[G[A_{g''}]] \\ &= \Delta^{-1}[A] \end{aligned} \tag{405}$$

Then

$$\mathbb{1} = \Delta[A] \int \mathscr{D}\vec{\omega}\, \delta[G[A_\omega]] \tag{406}$$

This little chappie can be inserted at will into the FPI!

Suppressing the source term, as before, we obtain

$$Z_{\rm YM} = \int \mathscr{D}A_\mu\, e^{iS[A]} = \int \mathscr{D}A_\mu\, \Delta[A_\mu] \int \mathscr{D}\vec{\omega}\, \delta[G[A_\omega]]e^{iS[A]} \tag{407}$$

We next perform a gauge transformation to turn As into A_ωs and then relabel the dummy variable of integration A_ω as A. After this finagle, which is just a simple consequence of gauge invariance, there is nothing left in the integrand depending on ω. Therefore, we can factor out the multiplicative constant $\int \mathscr{D}\omega$ – it will do absolutely no violence to the physics of our QFT if we pull it out front into the FPI normalization. This cannot affect correlation functions, as we proved earlier in the course, and the same goes for S-matrix elements. Therefore, we write

$$Z_{\rm YM} := \mathscr{N} \int \mathscr{D}A_\mu \Delta[A_\mu]\, \delta[G[A_\mu]]e^{iS[A]} \tag{408}$$

Our next step in this exposition is to actually *evaluate* $\Delta[A]$, starting from our definition of $\Delta^{-1}[A]$. To figure this out, we recall the infinitesimal form of gauge transformations

$$\begin{aligned} G^A[A_\omega] &= G^A[A] + \frac{\partial G^A}{\partial A^B_\mu}\, \delta A^B_\mu \\ &= G^A[A] + \frac{\partial G^A}{\partial A^B_\mu} \left(\delta^B_D \partial_\mu + f_{CD}{}^B A^C_\mu\right) \Delta\omega^D \end{aligned} \tag{409}$$

so that

$$G^A[A_\omega] = G^A[A] + \frac{\partial G^A}{\partial A^B_\mu} \left(D_\mu \Delta\omega\right)^B \tag{410}$$

Then, since the δ function enforces $G^A[A] = 0$, we obtain

$$\Delta^{-1}[A] = \int \mathscr{D}\vec{\omega}\ \delta\left[\frac{\partial G^A}{\partial A^B_\mu} D_\mu{}^B{}_C \Delta\omega^C\right] \tag{411}$$

Now let us denote the coefficient of $\Delta\omega$ in the argument of the delta function[al] by M:

$$\frac{\partial G^A}{\partial A^B_\mu} D_\mu{}^B{}_C \delta^D(x-y) \equiv M^A{}_C(x,y)\delta^D(x-y)$$

$$
\begin{aligned}
&\equiv \langle A, x|M|C, y\rangle \\
&\sim \frac{\delta G^A[A(x)]}{\delta\Delta\omega^C(y)}
\end{aligned} \tag{412}
$$

and then

$$\Delta^{-1}[A] = (\det M)^{-1} \tag{413}$$

We would also have expected this result just by eyeballing the coefficient of $\Delta\omega$ in the argument of the delta function[al] and generalizing from $\delta(ax) = (1/a)\delta(x)$.

To prove this result, suppose that our gauge-fixing function G^A has eigenvalues λ^i with eigenfunctions f^i. Then we have

$$
\begin{aligned}
\sum_{B,y} \langle A, x|M|B, y\rangle f^i_B(y) &= \lambda^i f^i_A(x) \\
\text{i.e. } \sum_{B,y} M_{ab}(x,y)\delta^D(x-y) f^i_B(y) &= \lambda^i f^i_A(x) \\
\text{so } \sum_B M_{AB}(x,x) f^i_B(x) &= \lambda^i f^i_A(x)
\end{aligned} \tag{414}
$$

Next, we expand our infinitesimal parameters in terms of the eigenfunctions of M:

$$\omega^C(y) = \sum_i \omega^i f^i_C(y) \tag{415}$$

Then

$$\sum_B M_{AB}(y,y)\omega^B(y) = \sum_i \omega^i \lambda^i f^i_A(y) \tag{416}$$

so that

$$\Delta^{-1}[A] = \int \mathscr{D}\omega^A \prod_A \delta\left[\sum_i \omega^i \lambda^i f^i_A(y)\right] \tag{417}$$

Now abbreviate

$$u_A \equiv \sum_i \omega^i \lambda^i f^i_A \tag{418}$$

then

$$
\begin{aligned}
\Delta^{-1}[A] &= \int \mathscr{D}u_A \frac{\partial(\omega^1, \omega^2, \cdots)}{\partial(u_1, u_2, \cdots)} \prod_A \delta(u_A) \\
&= \left.\frac{\partial(\omega^1, \omega^2, \cdots)}{\partial(u_1, u_2, \cdots)}\right|_{u=0} \\
&= (\lambda_1\lambda_2\cdots)^{-1} \left|f^i_A\right|^{-1} \\
&= \mathscr{N}' \frac{1}{\lambda_1\lambda_2\cdots} = \mathscr{N}' (\det M)^{-1}
\end{aligned} \tag{419}
$$

where in the last line we used the fact that any weightings by $f^i_A(x)$ will, when integrated (inside the FPI), just alter the overall field-independent normalization constant out front! Therefore, up to a physically irrelevant constant,

$$\Delta[A] = \det(M) \tag{420}$$

Suppose we chose for instance the Lorenz gauge

$$G^A[A] = \partial^\mu A_\mu^A \tag{421}$$

As we have seen already, the gauge condition $G^A = 0$ implies a $\delta[G[A]]$ in the FPI. Similarly, if

$$G^A[A] = \partial^\mu A_\mu^A + \beta^A(x) \tag{422}$$

where $\beta^A(x)$ is an arbitrary ordinary function, then the FPI integrand, as before, has a $\delta[G[A] - \beta]$. However, since this $\beta(x)$ is independent of the functional field A_μ, any integral over β will just contribute an irrelevant normalization constant out front. Inserting for instance

$$\exp\left(-\frac{i}{2\alpha}\int d^Dx\, \beta^A(x)\beta^A(x)\right) \tag{423}$$

gives

$$Z_{\text{YM}} = \mathscr{N}\int \mathscr{D}A_\mu \Delta[A] \exp\left(i\int d^Dx\left[\mathscr{L}_{\text{YM}} - \frac{1}{2\alpha}G[A]^2\right]\right) \tag{424}$$

We can now massage the $\Delta[A]$ by using our favourite Grassmann determinant trick:

$$\Delta[A] = \det(M) = \int \mathscr{D}\eta\mathscr{D}\bar{\eta} \exp\left(-i\bar{\eta}^A M_{AB}\eta^B\right) \tag{425}$$

Finally, we attain

$$Z_{\text{YM}}(\text{Lorentz}) = \mathscr{N}\int \mathscr{D}A_\mu\mathscr{D}\eta\mathscr{D}\bar{\eta} \exp\left[i\int\left(\mathscr{L}_{\text{YM}} - \frac{1}{2\alpha}G^2 - \bar{\eta}^A M_{AB}\eta^B\right)\right] \tag{426}$$

Therefore, the full Lagrangian for Yang-Mills gauge fields coupled to matter – including Fadeev-Popov ghosts – is

$$\mathscr{L} = \mathscr{L}_{\text{m}} + \mathscr{L}_{\text{YM}} + \mathscr{L}_{\text{gf}} + \mathscr{L}_{\text{gh}} \tag{427}$$

where

$$\mathscr{L}_{\text{m}} = i\bar{\psi}\not{D}\psi - m\bar{\psi}\psi = i\bar{\psi}_A\left(\delta^A{}_B\not{\partial} - ig\not{A}^C\left(T_C\right)^A{}_B\right)\psi^B - m\bar{\psi}_B\psi^B \tag{428}$$

with

$$\mathscr{L}_{\text{YM}} = -\frac{1}{4}\text{Tr}\left(F_A^{\mu\nu}F_{\mu\nu}^A\right) \tag{429}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu - ig[A_\mu, A_\nu]$. Also,

$$\mathscr{L}_{\text{gf}} = -\frac{1}{2\alpha}G^A G_A \tag{430}$$

and

$$\mathscr{L}_{\text{gh}} = \bar{\eta}_A\frac{\partial G^A}{\partial A_\mu^B}D_\mu^B{}_C\eta^C \tag{431}$$

This is the full Lagrangian for quantum Yang-Mills. It's all you need to compute everything there is to know about YM gauge fields minimally coupled to matter.

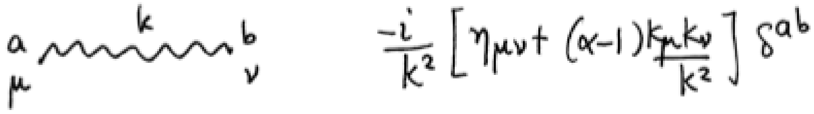


Figure 1: Propagator for Yang-Mills fields

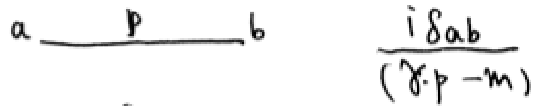


Figure 2: Propagator for matter fields

7.6 Lorenz gauge Feynman rules for Yang-Mills

Here, we have

$$G^A = \partial^\mu A^A_\mu \tag{432}$$

Also, recall that for A_μ in the Lie algebra for gauge group $\mathscr{G}$, A transforms under gauge transformations as

$$\delta A^A_\mu = f_{BC}{}^A A^B_\mu \omega^C + \partial_\mu \omega^A \tag{433}$$

Then

$$\delta G^A[A] = f_{BC}{}^A \partial^\mu \left(A^B_\mu \omega^C \right) + \Box \omega^A \tag{434}$$

so that our matrix in whose determinant we are interested is

$$M_{AB} = \frac{\delta G^A}{\delta \omega^B} = -f_{BC}{}^A \partial^\mu A^C_\mu - f^A_{BC} A^C_\mu \partial^\mu + \delta^{AB} \Box \tag{435}$$

Therefore, for the ghost piece of the FPI, we focus on

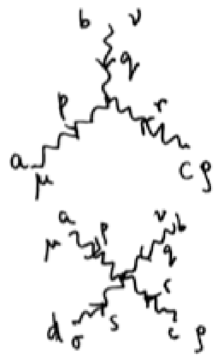
$$\begin{aligned} Z &\propto \int \mathscr{D}\eta \mathscr{D}\bar{\eta} \exp\left(-i \int d^D x \, \bar{\eta}^A \frac{\delta G^A}{\delta \omega^B} \eta^B \right) \\ &= \int \mathscr{D}\eta \mathscr{D}\bar{\eta} \exp\left(-i \int d^D x \left\{ \bar{\eta}^A \Box \eta^A - g f_{BC}{}^A \left[(\bar{\eta}^A \partial^\mu \eta^B) A^C_\mu - \partial^\mu A^C_\mu (\bar{\eta}^A \eta^B) \right] \right\} \right) \end{aligned} \tag{436}$$

In other words, in Lorenz gauge, F-P ghosts have Lagrangian

$$\mathscr{L}^{\text{Lorentz}}_{\text{gh}} = \bar{\eta}_A \Box \eta^A - g f_{BC}{}^A \left[\left(\bar{\eta}_A \partial^\mu \eta^B \right) A^C_\mu - \left(\partial^\mu A^B_\mu \right) \left(\bar{\eta}^A \eta^C \right) \right] \tag{437}$$

Propagators and vertices are shown in the following Fig.1-6.

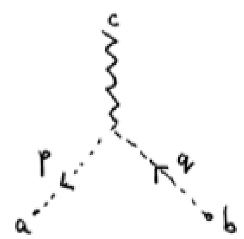
Figure 3: Propagator for ghost fields



$$-gf^{abc}[(r-q)_\mu\eta_{\nu\rho} + (p-r)_\nu\eta_{\mu\rho} + (q-p)_\rho\eta_{\mu\nu}]$$

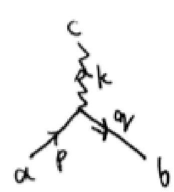
$$-g^2[f^{abe}f^{cde}(\eta_{\mu\rho}\eta_{\nu\sigma} - \eta_{\mu\sigma}\eta_{\nu\rho}) + f^{ace}f^{bed}(\eta_{\mu\sigma}\eta_{\rho\nu} - \eta_{\mu\nu}\eta_{\rho\sigma}) + f^{ade}f^{bce}(\eta_{\mu\nu}\eta_{\rho\sigma} - \eta_{\mu\rho}\eta_{\sigma\nu})$$

Figure 4: Yang-Mills cubic and quartic vertices



$$-gf^a{}_{bc}\,p_\mu$$

Figure 5: ghost-YM cubic vertex



$$-ig\gamma_\mu(T^c)_{ab}$$

Figure 6: matter-YM cubic vertex

7.7 BRST invariance and unitarity

You may worry that we inserted the $\mathscr{L}_{\text{gf}}$ piece of the Lagrangian rather arbitrarily. Let us now discuss a more mathematically sophisticated approach known as the BRST method, named for Becchi, Rouet, Stora, and Tyutin. As we will see, the form of the gauge-fixing Lagrangian is set by the form of the ghost Lagrangian, whose definition is, in turn, clear from the Fadeev-Popov procedure (its form is set by our friend the matrix M).

BRST noticed that if we choose the gauge transformation parameter $\Delta\omega^A(x)$ to be

$$\Delta\omega^A(x) = -\eta^A(x)\epsilon\,, \tag{438}$$

where $\eta^A(x)$ and ϵ are both Grassmann variables and ϵ is a constant, then the Yang-Mills action is invariant – provided that the field $\eta^A(x)$ transforms with a particular form. The full BRST transformation equations for matter, gauge, ghost, antighost and Lagrange multiplier fields $(\Psi, A_\mu, \eta, \bar{\eta}, \Pi)$ read:

$$\begin{aligned}
\delta_{\text{BRST}}\Psi(x) &= -i(T_B)\eta^B(x)\epsilon\Psi \\
\delta_{\text{BRST}}A^A_\mu(x) &= (D_\mu)^A{}_C\eta^C\epsilon \\
\delta_{\text{BRST}}\eta^A(x) &= -\frac{1}{2}f^A{}_{BC}\eta^B(x)\eta^C(x)\epsilon \\
\delta_{\text{BRST}}\bar{\eta}^A(x) &= \Pi^A(x)\epsilon \\
\delta_{\text{BRST}}\Pi(x) &= 0
\end{aligned} \tag{439}$$

Using the Jacobi identity satisfied by the structure constants, it is possible (although somewhat arduous algebraically) to show that the BRST transformation squares to zero:

$$\delta^2_{\text{BRST}} = 0 \tag{440}$$

Since the BRST operator generating these BRST symmetry transformations is nilpotent, any term added to the action which is the BRST-variation of something will automatically be BRST-invariant. Formally, we can write this as

$$\mathscr{L}_{\text{tot}} = \mathscr{L}_{\text{cl}} + \frac{\delta_{\text{BRST}}}{\delta\epsilon}\Xi \tag{441}$$

where Ξ is known as the gauge fermion. Choosing in particular

$$\Xi = \bar{\eta}^A\left(G^A + \frac{1}{2}\alpha\Pi^A\right) \tag{442}$$

gives

$$\mathscr{L}_{\text{tot}} = \mathscr{L}_{\text{YM}} + \Pi^A G^A + \frac{\alpha}{2}\Pi^A\Pi^A - \bar{\eta}^A\frac{\partial G^A}{\partial A^B_\mu}(D_\mu)^B{}_C\eta^C \tag{443}$$

Now, notice that $\Pi^A(x)$ is an *auxiliary* field: it does not propagate and so can be integrated out. After the integration over Π we obtain

$$\mathscr{L}_{\text{tot}} = \mathscr{L}_{\text{YM}} - \frac{1}{2\alpha}G^A G^A - \bar{\eta}^A\frac{\partial G^A}{\partial A^B_\mu}(D_\mu)^B{}_C\eta^C \tag{444}$$

which is what we wanted to show. In other words, the form of the gauge-fixing Lagrangian is intimately tied together with the form of the Fadeev-Popov ghost Lagrangian.

Let us switch for the remainder of this subsection to thinking in the canonical picture of QFT in order to glean some important physical facts. As Peskin and Schroeder point out, the structure of BRST transformations imply very nontrivial things about unitarity of scattering amplitudes. These all originate from having a nilpotent operator Q that commutes with the Hamiltonian H: $[H,Q]=0$. To see this, let $\mathscr{H}_1$ be the subspace of states $|\psi_1\rangle$ which are *not* annihilated by the BRST operator Q, and let $\mathscr{H}_2$ be the subspace of states $|\psi_2\rangle$ of the form $|\psi_2\rangle = Q|\psi_1\rangle$ for some $|\psi_1\rangle \in \mathscr{H}_1$. Also let $\mathscr{H}_0$ be the subspace of states $|\psi_0\rangle$ satisfying $Q|\psi_1\rangle = 0$ but that cannot be written as the Q of any state (the states which are BRST-closed but not BRST-exact). Having these three subspaces of the space of states $\mathscr{H}_0, \mathscr{H}_1, \mathscr{H}_2$ follows directly from having a nilpotent operator which commutes with the Hamiltonian. Incidentally, note that the subspace $\mathscr{H}_2$ is a rather weird subspace, as any two states in it have zero inner product. This fact also follows directly from the nilpotency of Q.

Consider one-particle states. By inspecting the form of the BRST transformations, we can see that in Lorenz gauge the BRST operator Q converts the forward component of the gauge field to the ghost field. By "forward" is meant polarization vectors $\epsilon^+_\mu(k)$ where

$$\epsilon^\pm_\mu(k) = \frac{1}{\sqrt{2}}\left(\frac{k^0}{|\vec{k}|}, \pm\frac{\vec{k}}{|\vec{k}|}\right) \tag{445}$$

These obey $(\epsilon^\pm)^2 = 0$ and $\epsilon^\pm \cdot \epsilon^\mp = 1$, as well as being orthogonal to the transverse polarizations: $\epsilon^\pm \cdot \epsilon^T_i = 0$. The transverse polarizations ϵ^T_i for $i=1,2$ obey $\epsilon^T_i \cdot \epsilon^{T*}_j = -\delta^i_j$. The completeness relation is $\eta_{\mu\nu} = \epsilon^+_\mu \epsilon^{-*}_\nu + \epsilon^-_\mu \epsilon^{+*}_\nu - \sum_i \epsilon^T_{i\mu}\epsilon^{T*}_{i\nu}$ in our mostly minus signature of spacetime. Q also converts the antighost into a quantum of the Π field which, through the classical equation of motion $\alpha\Pi^A = -G^A$ and the form of G^A in Lorenz gauge $G^A = \partial^\mu A^A_\mu$, must be a gauge boson which satisfies $k_\mu \epsilon^\mu(k) \neq 0$; these are the backwards-polarized gauge bosons.

So overall, among 1-particle states, $\mathscr{H}_1$ contains forward-polarized gauge bosons and antighosts, $\mathscr{H}_2$ contains ghosts and backward-polarized gauge bosons, while $\mathscr{H}_0$ contains only the physical transverse gauge bosons. More generally, it can be shown that asymptotic states containing ghosts, antighosts, or gauge bosons of unphysical polairzation always belong to $\mathscr{H}_1$ or $\mathscr{H}_2$, while asymptotic states in $\mathscr{H}_0$ contain only transversely polarized gauge bosons.

What does all this imply about unitarity? That is fairly straightforward to see. Let $|A;\perp\rangle$ denote an external state containing no ghosts or antighosts and only physical (transverse) polarizations of gauge bosons. We would like to show that the S-matrix for these guys is unitary, i.e. that

$$\sum_C \langle A;\perp|S^\dagger|C;\perp\rangle\langle C;\perp|S|B;\perp\rangle = \langle A;\perp|\mathbb{1}|B;\perp\rangle \tag{446}$$

Recall that $\mathscr{H}_0$ is the space of states $|\psi_0\rangle$ such that $Q|\psi_0\rangle = 0$ but which cannot be written as $Q|\lambda\rangle$ for *any* λ (i.e., the states which are BRST-closed but not BRST-exact). Also, we know that $[Q,H]=0$ so that any time-evolved state $S|\psi_0\rangle$ is also killed by the BRST operator Q: it takes the form

$$Q \cdot S|A;\perp\rangle = 0 \tag{447}$$

This implies that $S|A;\perp\rangle$ must be linear combinations of states in $\mathscr{H}_0$ (physical ones) and states in $\mathscr{H}_2$ (BRST-exact ones). But any two states in $\mathscr{H}_2$ have zero inner product with one another, and also $\langle\psi_2|\psi_0\rangle = 0$ (by definition of $\mathscr{H}_2$). So the inner product of any two states of that form must arise solely from the overlap of the components in $\mathscr{H}_0$. Therefore,

$$\langle A;\perp|S^\dagger S|B;\perp\rangle = \sum_C \langle A;\perp|S^\dagger|C;\perp\rangle\langle C;\perp|S|B;\perp\rangle \tag{448}$$

and so not only is the full S-matrix unitary but its restriction to the subspace $\mathscr{H}_0$ is also unitary. This is the physical reason why Feynman diagrams producing pairs of gauge bosons with unphysical polarizations must always be cancelled by diagrams producing ghosts. (c.f. Cutkosky rules to be discussed during 1-loop renormalization of QED.) Neato!

BRST symmetry of the action (with Yang-Mills, gauge-fixing and Fadeev-Popov terms) implies identities that are satisfied by the generating functional Z. These identities can be used to derive identities for the connected generating functional W or for its Legendre-transformed friend Γ. The resulting identities are usually referred to as Ward Identities. For the case of $U(1)$ we will discuss the Ward identities soon, when we get to 1-loop QED. The non-Abelian counterparts are known as the Slavnov-Taylor identities and we will not have time to develop them in this course.

8 Renormalization and Quartic Scalar Field Theory

8.1 Length Scales

Nonlinear equations govern the long-wavelength dynamics of fluids. This hydrodynamic description works great at long wavelengths, but is obviously a very poor one on molecular scales, where graininess matters immensely. At long distances, the fact of complex atomic and molecular quantum dynamics going on at short distances is physically irrelevant: fluid mechanics does just fine as a classical theory of fluids. We do not need to know about quantum mechanics just to compute long-wavelength physics.

Therefore, it is important to match the effective field theory – the working Lagrangian – to the typical length scale in the problem. For instance, we would presumably rather use density functional theory than the Navier-Stokes equation for investigating the properties of water on nanometre scales. Equally, it would be a total waste of time to start from the equations of quantum string theory to describe the orbit of Jupiter around our Sun. Newtonian physics was accurate enough to land a human on the Moon, after all!

Separation of scales, in which short-wavelength physics does not affect long-wavelength physics, is known as *universality*. It turns out to be a much more general phenomenon, forming a cornerstone of the modern Wilsonian approach to understanding ultraviolet properties of loop Feynman graphs in QFTs. It is rather involved to demonstrate universality mathematically. The basic physics idea behind it is to parametrize the effect of all possible UV modifications of the QFT at higher energy (and shorter distance) in terms of all possible interactions between the effective low-energy degrees of freedom. For well-behaved renormalizable QFTs, this will result in quantum shifts of masses and couplings at loop level. For non-renormalizable QFTs, trying to renormalize at one more loop order generates new terms in the low-energy effective Lagrangian, ad infinitum, giving an infinite mess of infinities.

8.2 Cutoffs

As you calculated in HW2, loop Feynman graphs typically diverge. The physical reason why loop integrals blow up is the assumption that loop momenta k may vary from 0 to ∞. Nobody yet knows exactly how spacetime works at short distance, especially at lengths below the Planck scale where gravity becomes quantum mechanical. It is therefore preposterous to claim that any given QFT will be valid all the way to $k = \infty$. This is one of the main reasons why physicists do not lose as much sleep as mathematicians in thinking about QFT at loop level: we know that some other description of the physics must take over at very short wavelengths.

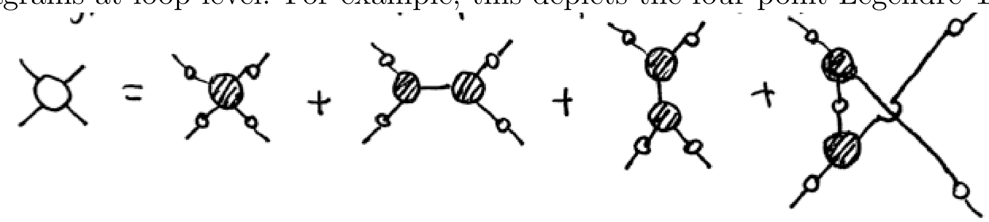
We might more reasonably expect any given QFT to work up to some upper energy limit, usually denoted as Λ, but not above that scale. We can think of Λ as an *ultraviolet cutoff*, an energy scale beyond which detailed knowledge of short-wavelength physics is essential to make physical predictions. At momenta much lower than the UV cutoff, the effective Lagrangian encoding the physics of only the low-energy modes will be insensitive to the UV physics – *except* through the parameters of that theory such as mass and coupling strength. As we will see next week in some detail, if Λ is changed, then the only effect in the low-energy effective Lagrangian is to alter the masses and couplings – a story which may include quantum generation of couplings that were absent in the classical theory. The low-energy values of all masses and couplings, such as $\alpha_{\rm EM}(0) \sim 1/137.036$, must of course then match onto experiment.

The physical process of making sense of mathematical infinities in loop Feynman graphs is known as renormalization. Many different schemes are available for renormalization, and they all depend explicitly on the details of how mathematical infinities are regulated. One renormalization scheme involves latticizing space (or Euclidean spacetime), with a spacing $1/\Lambda$. This lattice scheme is actually the only one known to be well-defined for gauge field theories non-perturbatively. Perturbatively speaking, there are two useful Lorentz-invariant schemes in common use today: Pauli-Villars, and (by far the most popular one) Dimensional Regularization. We will use the latter. In this scheme, we will continue the dimension of spacetime away from $D = 4$ and inspect infinities arising in the limit $D - 4 \to 0$. This is a formal manipulation only and should not be interpreted in terms of fractal dimension. Indeed, mathematically, DR does not correspond to a positive measure. UV divergences will show up as *poles* at discrete values of D.

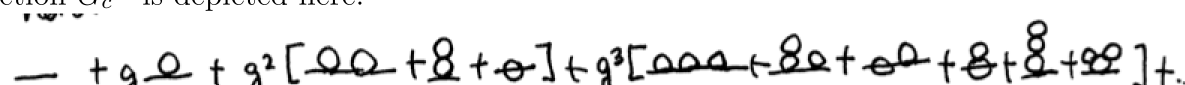
Aspects of the discussion in the rest of this chapter and the next are taken from §9 of the introductory QFT textbook by Lewis Ryder.

8.3 Focus: 1PI Diagrams in ϕ^4

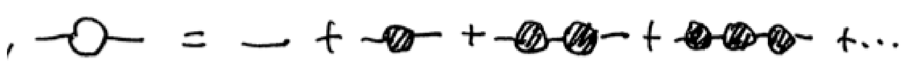
Significantly earlier in this course, we learned that only connected Feynman graphs contribute to correlation functions of interest: disconnected ones cancel out between the numerator and denominator in our formula for correlation functions in path integral quantization. We also learned that only 1PI diagrams are needed in order to construct connected Feynman diagrams at loop level. For example, this depicts the four-point Legendre Trees formula:



For propagator corrections there is an analogous separation into 1PI 'nuggets' that we have not yet discussed. Consider all possible loop corrections to the Feynman propagator in scalar field theory with quartic potential. This structure for the connected two-point function $G_c^{(2)}$ is depicted here:-



Notice how several of the graphs are not 1PI. Indeed, as you should check explicitly, the complete animal, known as the *dressed propagator*, can be written in terms of the 1PI *proper self-energy* as follows:



In this figure, the complete propagator is represented by the open-circle piece, while the 1PI contribution has the shaded interior. From this graphical expansion and basic knowledge of geometric series, we can see quickly that

$$G_c^{(2)}(p) = \frac{i}{p^2 - m^2 - \Sigma(p)} \tag{449}$$

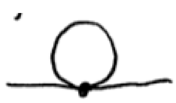
where $\Sigma(p)$ represents the proper self-energy. Accordingly, the physical mass of the ϕ field will be

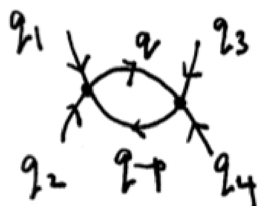
$$m^2_{\text{phys}} = m^2 + \Sigma(p) \tag{450}$$

Note that if $\Sigma(p)$ blows up in the UV, we can compensate for this in matching to physical parameters by allowing the *bare* mass in the Lagrangian to be infinite, in just such a way as to cancel the infinity from the physical quantity. We will see how this works much more explicitly soon.

8.4 Divergences in ϕ^4 in $D = 4$

What kind of divergences do we face in attempting to understand loop level physics for general QFTs? Let us begin this analysis with examples from ϕ^4 at one loop. The 1-loop 1PI diagrams which renormalize mass and coupling are, respectively,



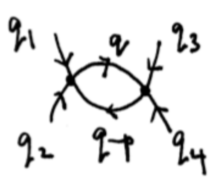


Let us estimate how these integrals will blow up in the UV. For the first diagram, we have at one loop

$$= \quad g \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 - m^2 + i\epsilon}$$
$$\to \Lambda^2 \text{ as } q \to \infty \text{ in } D = 4 \tag{451}$$

where Λ is the (momentum) UV cutoff. In other words, the propagator correction is quadratically divergent in four spacetime dimensions.

How about the coupling correction, also at one loop? There, we have



$$= \quad \frac{g^2}{2} \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 - m^2 + i\epsilon)} \frac{1}{[(q - q_1 - q_2)^2 - m^2 + i\epsilon]}$$
$$\to \log(\Lambda) \text{ as } q \to \infty \text{ in } D = 4. \tag{452}$$

In other words, the coupling correction at one loop is logarithmically divergent for ϕ^4 theory in $D = 4$.

8.5 Divergences in General

We can be more systematic about this. Indeed, let us consider a general Feynman graph for ϕ^k field theory with V vertices, E external lines, I internal lines and L loops. in D dimensions, this gives the *superficial degree of divergence* $\mathscr{D}$ as

$$\mathscr{D} = DL - 2I \tag{453}$$

The first term is clearly just emanating from the integration measures for loop momenta. The second part encodes the fact that the number of propagators involved in the loop integration is driven by the number of internal lines.

For an L-loop diagram, with E external lines and I internal ones, momentum conservation is required at each of the V vertices. Overall momentum conservation and momentum conservation at each vertex yields $(V-1)$ relations amongst momenta. Therefore, the number of independent momenta is

$$L = I - V + 1 \tag{454}$$

i.e. the number of loops. So

$$\mathscr{D} = (D-2)I + D(1-V) \tag{455}$$

How does this relate to E? Well, every vertex can have either (i) an external leg connected to one vertex, or (ii) an internal leg, which must by its nature be connected to a second vertex as well as the first. Therefore,

$$kV = E + 2I \tag{456}$$

Eliminating L and I in favour of D, E, V gives

$$\mathscr{D} = D - \left(\frac{D}{2} - 1\right) E + \left[\frac{(D-2)k}{2} - D\right] V \tag{457}$$

For quartic scalar field theory in 4D, $k = 4$, which gives $\mathscr{D} = 4 - E$, independent of V. If instead we had ϕ^k field theory in 4D for $k > 4$, then $\mathscr{D}$ would grow with the number of vertices V, meaning that we will sprout more and more divergent integrals as we go to higher and higher loop order. These are known as *non-renormalizable* or *irrelevant* interactions. As you can convince yourself by doing simple power counting in the effective Lagrangian for the kinetic vs interaction terms, these non-renormalizable cases are also precisely the cases for which the coupling of that interaction must have negative mass dimension. Conversely, an interaction with positive mass dimension will have fewer divergent integrals as we go to higher order in perturbation theory. These are known as *super-renormalizable* or *relevant* interactions. Interactions with zero mass dimension are known as *renormalizable* or *marginal*. In 3D, a ϕ^6 becomes marginal while ϕ^4 is relevant. In 2D, any ϕ^n coupling is relevant.

The structure of renormalization is quite different for irrelevant as compared to marginal or relevant perturbations. For irrelevant ones, we will generate every interaction consistent with the symmetries of the problem if we go to high enough order in perturbation theory. So irrelevant interactions have no cutoff-independent physical meaning. By contrast, for marginal or relevant perturbations, the subtraction process stops with a finite number of operators in the effective Lagrangian.

8.6 Weinberg's Theorem

The superficial degree of divergence is not all there is to know about a Feynman diagram in dimensional analysis. It also turns out to be necessary to consider the degree of divergence of each possible subgraph, obtained by holding one of the loop momenta fixed.

We will state but not prove the following immensely powerful theorem:

> A Feynman diagram converges if its degree of divergence $\mathscr{D}$, together with the degree of divergence of all its subgraphs, is negative.

Basically, what this says is that if you encounter a subgraph in your graph at some loop order (which will necessarily have fewer loops than the whole), then as long as you properly renormalized that subgraph at lower order in perturbation theory then you are home and hosed.

The difficulty in proving renormalizability of a QFT to all orders in perturbation theory is the phenomenon of *overlapping divergences*. To illustrate this, consider the "setting sun" diagram in $D = 4$ ϕ^4 theory in the high-k limit where mass effects are subleading:

$$\propto \int \frac{d^4k_1\, d^4k_2}{k_1^2 k_2^2 (p - k_1 - k_2)^2} \tag{458}$$

This diagram has overall degree of divergence $\mathscr{D} = 8 - 6 = 2$. However, if we hold k_1 fixed, the integral over k_2 gives $\mathscr{D} = 4 - 4 = 0$ i.e. a logarithmic divergence. The same holds for $1 \leftrightarrow 2$. So it is impossible to separate out the divergences in k_1 from those in k_2.

In gauge theories such as QED, gauge invariance prohibits overlapping divergences, which is completely awesome.

8.7 Dimensional Regularization and the Propagator Correction

One of the sexiest things about Dimensional Regularization (DR) is that it is a renormalization scheme which preserves Poincaré invariance and nonAbelian gauge invariance.

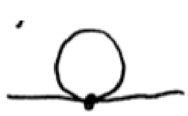
Since D will be continued away from $D = 4$, we have to be careful about counting mass dimensions of interaction terms. We work out what the mass dimension of a field is by insisting that its quadratic kinetic term in the effective Lagrangian remains unmolested, pushing the dependence on any mass scale into the interaction. We have

$$\mathscr{L}_0 = \frac{1}{2}(\partial\phi)^2 - \frac{1}{2}m^2\phi^2 - \frac{1}{4!}g\mu^\varepsilon\phi^4 \tag{459}$$

where

$$\varepsilon \equiv 4 - D \tag{460}$$

and μ is a mass scale (theory parameter) and g is dimensionless. Then



$$= \frac{g\mu^\varepsilon}{2}\int \frac{d^Dq}{(2\pi)^D}\frac{1}{q^2 - m^2 + i\epsilon} \tag{461}$$

Note that in the above, ϵ is the small parameter reminding us to use the Feynman propagator, while ε is the amount we continue away from $D = 4$ in order to regularize our infinite loop integrals. They are not at all related!

The process of evaluating loop integrals like the above, and more complex ones, can be done by recruiting a few main conceptual steps. First, we use Feynman parameters (p of them for $p + 1$ propagator denominators) to turn the denominator from a product of propagators into a power of a single propagator-like denominator. Second, we get rid of any linear terms in a quadratic denominator by shifting the loop momentum and completing the square. Third, we Wick rotate to Euclidean signature to avoid getting confused about square roots of timelike momenta. Fourth, we use spherical polar coordinates in momentum space to actually perform the integrals. The upshot is a set of rules about momentum integrals that you can find in Appendix A.4 in Peskin and Schroeder, starting on p.805. The main one that we need here is

$$\int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 - \Delta)^n} = \frac{(-1)^n\, i}{(4\pi)^{D/2}} \frac{\Gamma(n - D/2)}{\Gamma(n)} \frac{1}{\Delta^{n-D/2}} \,. \tag{462}$$

Gamma functions have poles at zero and negative integers. A useful math fact is

$$\Gamma(-n + \varepsilon) = \frac{(-1)^n}{n!} \left[\frac{1}{\varepsilon} + \psi_1(n+1) + \mathscr{O}(\varepsilon) \right] , \tag{463}$$

where

$$\psi_1(n+1) = 1 + \frac{1}{2} + \ldots + \frac{1}{n} - \gamma \,, \tag{464}$$

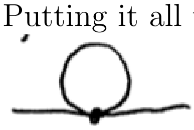
and γ is the Euler-Mascheroni constant. Therefore,

$$\Gamma\left(1 - \frac{D}{2}\right) = \Gamma\left(-1 + \frac{\varepsilon}{2}\right) = -\frac{2}{\varepsilon} - 1 + \gamma + \mathscr{O}(\varepsilon) \,. \tag{465}$$

Finally, recall that for $a \in \mathbb{R}$ and small ε

$$a^\varepsilon = 1 + \varepsilon \ln(a) + \mathscr{O}(\varepsilon) \,. \tag{466}$$

Putting it all together, we obtain for the propagator renormalization



$$= \left(\frac{igm^2}{16\pi^2}\right) \frac{1}{\varepsilon} + \frac{igm^2}{32\pi^2} \left[1 - \gamma + \ln\left(\frac{4\pi\mu^2}{m^2}\right) \right] + \mathscr{O}(\varepsilon) \,. \tag{467}$$

This is supposed to depend on μ.

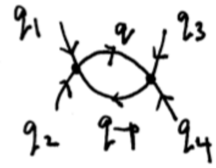
What does this say about the $\mathscr{O}(g)$ 1-loop contribution to the proper self-energy $\Sigma(p)$ introduced earlier? We have from our dimensional regularization adventure that

$$\Sigma = -\frac{gm^2}{16\pi^2 \varepsilon} + (\text{finite}) \; + \mathscr{O}(g^2) \,, \tag{468}$$

and therefore

$$\Gamma^{(2)}(p) = p^2 - m^2 \left(1 - \frac{gm^2}{16\pi^2 \varepsilon}\right) \,. \tag{469}$$

8.8 Vertex Correction



$$= \frac{g^2}{2}(\mu^2)^{\varepsilon} \int \frac{d^D q}{(2\pi)^D} \frac{1}{(q^2 - m^2 + i\epsilon)} \frac{1}{[(q-p)^2 - m^2 + i\epsilon]} \tag{470}$$

We could use a simple old standby *partial fractions*

$$\frac{1}{ab} = \frac{1}{(b-a)}\left(\frac{1}{a} - \frac{1}{b}\right) \tag{471}$$

Even more useful is *Feynman's formula*

$$\frac{1}{ab} = \int_0^1 \frac{dz}{[az + b(1-z)]^2} \tag{472}$$

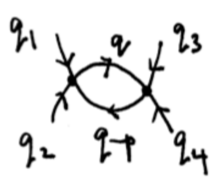
where $a, b \in \mathbb{C}$ and z is known as the *Feynman parameter*. For our 1-loop vertex correction graph, we use this to write

$$\frac{1}{(q^2 - m^2)[(q-p)^2 - m^2]} = \int_0^1 \frac{dz}{[q^2 - m^2 - 2qp(1-z) + p^2(1-z)^2]^2} \tag{473}$$

We now shift our loop momentum to get rid of the qp term:

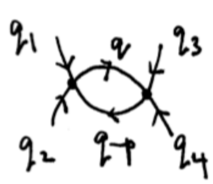
$$q' = q + p(1-z) \tag{474}$$

Using this, our integrand denominator becomes a perfect square. Relabelling $q' := q$ to save on worldwide usage of ', we obtain



$$\begin{aligned} &= \frac{1}{2} g^2 \mu^{2\varepsilon} \int_0^1 dz \int \frac{d^D q}{(2\pi)^D} \frac{1}{[q^2 - m^2 + p^2 z(1-z)]^2} \\ &= \frac{1}{2} g^2 \mu^{2\varepsilon} \frac{1}{\sqrt{4\pi}^D} \frac{\Gamma(2 - \frac{D}{2})}{\Gamma(2)} \int_0^1 dz \left[p^2 z(1-z) - m^2\right]^{D/2-2} \\ &= \frac{ig^2}{32\pi^2} \mu^{2\varepsilon} \Gamma(2 - \tfrac{D}{2}) \int_0^1 dz \left[\frac{p^2 z(1-z) - m^2}{4\pi\mu^2}\right]^{D/2-2} \end{aligned} \tag{475}$$

As $D \to 4$, $\Gamma(2 - \frac{D}{2}) \to 2/\varepsilon - \gamma + \mathscr{O}(\varepsilon)$ so that



$$\begin{aligned} &= \frac{ig^2}{32\pi^2} \mu^{2\varepsilon} \left(\frac{2}{\varepsilon} - \gamma + \mathscr{O}(\varepsilon)\right) \times \\ &\quad \times \left\{1 - \frac{\varepsilon}{2} \int_0^1 dz \ln\left[\frac{p^2 z(1-z) - m^2}{4\pi\mu^2}\right]\right\} \end{aligned} \tag{476}$$

$$\begin{aligned} &= \frac{ig^2}{16\pi^2} \mu^{2\varepsilon} \frac{1}{\varepsilon} - \frac{ig^2}{32\pi^2} \mu^{2\varepsilon} \times \\ &\quad \times \left\{\gamma + \int_0^1 dz \ln\left[\frac{p^2 z(1-z) - m^2}{4\pi\mu^2}\right]\right\} \end{aligned} \tag{477}$$

The Mandelstam variables are

$$\begin{aligned} s &= (q_1 + q_2)^2 \\ t &= (q_1 + q_3)^2 \\ u &= (q_1 + q_4)^2 \end{aligned} \tag{478}$$

The integrand in our above expression is

$$F(s, m, \mu) = \int_0^1 dz \ln \left\{ \frac{sz(1-z) - m^2}{4\pi\mu^2} \right\} \tag{479}$$

is a function only of $(q_1 + q_2)^2$ and m^2, as well as the renormalization scale μ . Note that the argument of the logarithm develops a branch cut when $(q_1 + q_2)^2 = 4m^2$. This is the harbinger of on-shell pair production.

In order to find the full 1-loop vertex correction, we amputate the external legs,

$$\Gamma^{(4)}(p_1, \ldots, p_4) = G^{(2)}(p_1)^{-1} \ldots G^{(2)}(p_4)^{-1} G^{(4)}(p_1, \ldots, p_4) \tag{480}$$

and put the s-channel diagram we just computed together with the t-channel and u-channel analogs. We obtain

$$\Gamma^{(4)}(p_i) = -ig\mu^\varepsilon \left(1 - \frac{3g}{16\pi^2\varepsilon}\right) - \left[\frac{ig^2\mu^\varepsilon}{32\pi^2} \{3\gamma + F(s, m, \mu) + F(t, m, \mu) + F(u, m, \mu)\}\right] \tag{481}$$

9 Callan-Symanzik equation and Wilsonian Renormalization Group

9.1 Counterterms

Previously, we discovered that a 1-loop divergence in a Feynman graph for ϕ^4 scalar field theory appears in dimensional regularization as a $1/\varepsilon$ pole as $\varepsilon = (4-D) \to 0$. The physical antidote to this unpleasant mathematical fact is to add to the Lagrangian an equal and opposite counterterm set to kill the pole in $1/\varepsilon$. These counterterms are denoted by Feynman graphs with one fewer loop order and with a circled-x:

Why does this idea work? For our scalar QFT, let us write counterterms as[10]

$$\mathscr{L}_{\rm CT} = \delta_Z \frac{1}{2} (\partial\phi)^2 - \delta_m \frac{m^2}{2}\phi^2 - \delta_g \frac{\mu^\varepsilon}{4!}\phi^4 \,, \tag{482}$$

Then note that we can write the full Lagrangian of our theory, known as the **bare Lagrangian**, as a sum

$$\mathscr{L}_{\rm B} = \mathscr{L}_0 + \mathscr{L}_{\rm CT} = \frac{1}{2}(1+\delta_Z)(\partial\phi)^2 - \frac{1}{2}m^2(1+\delta_m)\phi^2 - (1+\delta_g)\frac{g\mu^\varepsilon}{4!}\phi^4 \,. \tag{483}$$

This can be put in a more familiar looking form

$$\mathscr{L}_B = \tfrac{1}{2}(\partial\phi_B)^2 - \tfrac{1}{2}m_B^2\phi_B^2 - \frac{g_B}{4!}\phi_B^4 \tag{484}$$

by defining **renormalized fields and couplings** ϕ, m, g in terms of the **bare fields and couplings** ϕ_B, m_B, g_B via

$$\phi_B = \sqrt{Z_\phi}\,\phi \,, \quad Z_\phi = (1+\delta_Z) \tag{485}$$

$$m_B = Z_m m \,, \quad Z_m^2 = \frac{(1+\delta_m)}{(1+\delta_Z)} \tag{486}$$

$$g_B = Z_g g \mu^\varepsilon \,, \quad Z_g = \frac{(1+\delta_g)}{(1+\delta_Z)^2} \tag{487}$$

A QFT is said to be **renormalizable** if $\mathscr{L}_{\rm CT}$ is composed of terms of the same type as were originally present in the Lagrangian, and we see that very explicitly here.

Note: the *Minimal Subtraction (MS)* renormalization scheme is defined by keeping only the pole terms in the relation between renormalized and bare fields and parameters. The $\overline{\rm MS}$ scheme is a modification of MS designed to get rid of a bunch of transcendental numbers, allowing just enough of a finite part to get rid of the transcendentals.

[10] δ_Z is zero here only to one-loop. Note that our conventions for δ_m and δ_g differ from those in P&S.

9.2 Callan-Symanzik Equation

Recall that we had to introduce an energy scale μ in order to take care of the engineering dimensions of the coupling for our scalar field theory in D dimensions. Let us consider the logarithmic derivative $\mu\partial/\partial\mu$ of pertinent vertex functions. An unrenormalized vertex has no dependence on renormalization scale μ because it has no knowledge of the renormalization procedure. Using this and the conventions above for wavefunction renormalization, we see that the renormalized vertex obeys

$$\mu\frac{\partial}{\partial\mu}\left\{Z_\phi^{-n/2}\Gamma^{(n)}(p_i;g,m,\mu)\right\}=0\,. \tag{488}$$

Therefore, the renormalized vertex obeys the equation

$$\left[-n\mu\frac{\partial}{\partial\mu}\ln\sqrt{Z_\phi}+\mu\frac{\partial}{\partial\mu}+\mu\frac{\partial g}{\partial\mu}\frac{\partial}{\partial g}+\mu\frac{\partial m}{\partial\mu}\frac{\partial}{\partial m}\right]\Gamma^{(n)}(p_i;g,m,\mu)=0 \tag{489}$$

It is traditional to define

$$\begin{aligned}\beta(g)&=\mu\frac{\partial}{\partial\mu}g\\ \gamma(g)&=\mu\frac{\partial}{\partial\mu}\ln\sqrt{Z_\phi}\\ m\gamma_m(g)&=\mu\frac{\partial}{\partial\mu}m\end{aligned} \tag{490}$$

from which it follows immediately that

$$\left[\mu\frac{\partial}{\partial\mu}+\beta(g)\frac{\partial}{\partial g}-n\gamma(g)+m\gamma_m(g)\frac{\partial}{\partial m}\right]\Gamma^{(n)}(p_i;g,m,\mu)=0 \tag{491}$$

This is known as the Callan-Symanzik equation and dictates how the renormalized vertex *runs* with renormalization [energy] scale μ. $\beta(g)$ is known as the **beta function** of the coupling "constant" g while γ are the **anomalous dimensions**.

We can write down an alternate equation expressing the invariance of $\Gamma^{(n)}$ under scale transformations. Suppose that we scale by a factor t:

$$p\to tp;\,,\quad m\to tm;\,,\quad \mu\to t\mu\,. \tag{492}$$

Now, $\Gamma^{(n)}$ has mass dimension d, which in $D=4-\varepsilon$ dimensions is

$$d=D+n\left(1-\frac{D}{2}\right)=(4-n)+\varepsilon\left(\frac{n}{2}-1\right) \tag{493}$$

Then we have

$$\Gamma^{(n)}(tp_i;g,m,\mu)=t^d\Gamma^{(n)}(p_i;g,t^{-1}m,t^{-1}\mu) \tag{494}$$

So

$$\left(t\frac{\partial}{\partial t}+m\frac{\partial}{\partial m}+\mu\frac{\partial}{\partial\mu}-d\right)\Gamma^{(n)}=0 \tag{495}$$

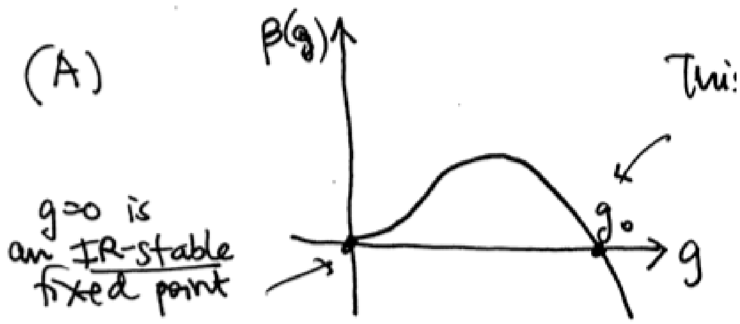
Eliminating $\mu \partial \Gamma / \partial \mu$ using the Callan-Symanzik equation gives

$$\left[-t\frac{\partial}{\partial t} + \beta\frac{\partial}{\partial g} - n\gamma(g) + m\left(\gamma_m(g) - 1\right)\frac{\partial}{\partial m} + d\right]\Gamma^{(n)}(tp; g, m, \mu) = 0 \tag{496}$$

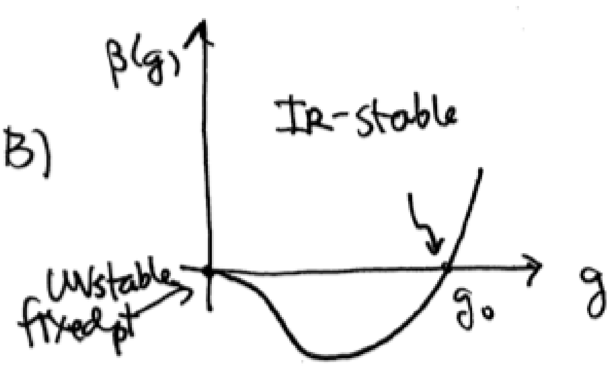
This equation gives the effect of scaling up momenta by a factor t. Notice how if the beta function and anomalous dimensions happen to vanish then the equation simply encodes the fact that Γ has canonical dimension d, as you would have expected from plain old ordinary dimensional analysis. Therefore, we see that *quantum corrections will change the scaling dimensions of vertices in the quantum effective action.* This is a very important aspect of loop level physics in QFT.

9.3 Fixed Points

A *fixed point* is a place at which the beta function vanishes, which may happen at a finite value of the coupling. Here are a couple of scenarios (A and B) for how "RG flow" might work in a QFT with coupling g.



Consider scenario (A) with a beta function as illustrated in the above figure. At $g \to g_0^+$, $\beta(g)$ is negative which is the slope for running of g. So the coupling is driven back to g_0 as μ increases. Similarly for $g \to g_0^-$. A fixed point like this is known as a UV-stable fixed point and obeys $g(\mu = \infty) = g_0$.



In scenario (B), by contrast, the beta function has opposite sign. So a fixed point at nonzero coupling will be a stable fixed point in the direction of *de*creasing μ, i.e., the IR. This case obeys $g(\mu = 0) = g_0$.

How does it work for ϕ^4?

Let us pretend that our 1-loop coupling renormalization is a good guide to the asymptotic behaviour. Ignoring finite corrections, at lowest order nontrivial order we had

$$g_B = (g\mu^\varepsilon)\left(1 + \frac{3g}{16\pi^2\varepsilon}\right) \tag{497}$$

Now, we know that the coupling g_B in the bare Lagrangian does not depend on μ. We can use this to find the logarithmic derivative of g and then the beta function, which is its limit as $\varepsilon \to 0$,

$$\beta(g) = \frac{3g^2}{16\pi^2} > 0 \tag{498}$$

At lowest nontrivial loop order (1-loop), this has solution

$$g = \frac{g_0}{[1 - ag_0 \ln(\mu/\mu_0)]} \tag{499}$$

where g_0 is a constant and

$$a = \frac{3}{16\pi^2} \tag{500}$$

Notice that the quartic self-interaction coupling increases as the renormalization scale μ increases. In other words, this theory is not asymptotically free. Vice versa, the beta function says that the coupling must get driven to zero in the IR, making ϕ^4 *trivial.*

Gauge field theories are qualitatively different. As we will see soon, they obey symmetry-driven identities known as Ward (Abelian) or Slavnov-Taylor (non-Abelian) identities. These ensure that the theory is renormalizable to all orders in perturbation theory, even in a Higgs phase. Non-Abelian gauge theories *are*, as you will see in your final project, asymptotically free.

9.4 Wilsonian RG and UV cutoffs

There is a more physically solid way to understand how changing the renormalization scale results in (a) running of couplings, and (b) generation of all terms in the quantum action allowed by symmetry whether they were present classically or not. This is the 1970s advance of Kenneth Wilson and remains one of the gems in the modern quantum physics literature. The description we outline here is taken from §12.1 of Peskin and Schroeder.

A simple prescription for avoiding loop divergences in QFTs would presumably be to avoid integrating over the high momenta which give rise to those divergences. For instance, we could nominate to integrate only over $\phi(k)$ such that $|k| \leq \Lambda$ and impose $\phi(k) = 0$ for $|k| > \Lambda$. This naïve prescription is fairly close to the mark, but we do have to be careful when putting in a hard UV cutoff like this: we live in Minkowski spacetime, which does not have a positive definite metric. The physically correct route is to use Euclidean space plus Wick rotation as our guide to the correct way to apply the UV cutoff in Lorentzian signature.

Let us now write the functional integral for quartic scalar field theory. For simplicity of exposition we shall set $J = 0$. Then we define our Feynman path integral as

$$Z = \int [\mathscr{D}\Phi]_\Lambda \exp\left(-\int d^Dx \left[\tfrac{1}{2}(\partial\Phi)^2 + \tfrac{1}{2}m^2\Phi^2 + \frac{\lambda}{4!}\Phi^4\right]\right) \tag{501}$$

where

$$[\mathscr{D}\Phi]_\Lambda = \prod_{|k|<\Lambda} d\Phi(k) \tag{502}$$

Divide $\Phi(k)$ into two groups: high-momentum and low-momentum modes

$$\begin{aligned} &\text{high momentum} : b\Lambda \le |k| \le \Lambda\,, \quad b \in (0,1) \\ &\text{low momentum} : |k| < b\Lambda \end{aligned} \tag{503}$$

Define

$$\hat{\phi}(k) = \begin{cases} \Phi(k)\,, & b\Lambda \le |k| < \Lambda \\ 0 & \text{otherwise} \end{cases} \tag{504}$$

Also define a new

$$\phi(k) = \begin{cases} \Phi(k)\,, & |k| < b\Lambda \\ 0 & \text{otherwise} \end{cases} \tag{505}$$

So we have

$$\Phi(k) = \hat{\phi}(k) + \phi(k) \tag{506}$$

Therefore

$$\begin{aligned} Z &= \int \mathscr{D}\phi \int \mathscr{D}\hat{\phi} \exp\left(-\int d^Dx \left[\tfrac{1}{2}(\partial\hat{\phi} + \partial\phi)^2 + \tfrac{1}{2}m^2\left(\phi + \hat{\phi}\right)^2 + \frac{\lambda}{4!}(\phi + \hat{\phi})^4\right]\right) \\ &= \int \mathscr{D}\phi \exp\left(-\int d^Dx \mathscr{L}[\phi]\right) \times \\ &\quad \times \int \mathscr{D}\hat{\phi} \exp\left(-\int d^Dx \left[\tfrac{1}{2}(\partial\hat{\phi})^2 + \tfrac{1}{2}m^2\hat{\phi}^2 + \right.\right. \\ &\qquad \left.\left. + \lambda\left\{\frac{1}{6}\phi^3\hat{\phi} + \frac{1}{4}\phi^2\hat{\phi}^2 + \frac{1}{6}\phi\hat{\phi}^3 + \frac{1}{4!}\hat{\phi}^4\right\}\right]\right) \end{aligned} \tag{507}$$

Note: any pieces proportional to $\phi\hat{\phi}$ vanish, by orthogonality of distinct Fourier modes.

Notice what just happened! The UV-cut-off functional integral has factored. The first factor, as you can see above, is completely independent of $\hat{\phi}$.

How would we perform the path integration over the $\hat{\phi}$s? We would like to end up with an expression only in terms of the lower-momentum stuff, which we would then interpret as our theory at the lower cutoff scale $b\Lambda$. We will mostly be interested in physical systems with $m^2 \ll \Lambda^2$ so that we can treat m^2 and λ as perturbations. The case of massless fields is somewhat trickier, as IR singularities (expressing long-range behaviour) can afflict the physics as well. For now, we take any masses to be finite but well below the UV cutoff scale.

At leading order, the portion of the Lagrangian involving $\hat{\phi}$ is

$$\int \mathscr{L}_0 = \int_{b\Lambda \le |k| < \Lambda} \frac{d^Dk}{(2\pi)^D} \hat{\phi}^*(k) k^2 \hat{\phi}(k) \tag{508}$$

Denoting Wick contractions by horizontal overbraces, we have for the two-point function giving the propagator

$$\overbrace{\hat{\phi}(k)\hat{\phi}(p)} = \frac{\int \mathscr{D}\hat{\phi}\, e^{-\int \mathscr{L}_0} \hat{\phi}(k)\hat{\phi}(p)}{\int \mathscr{D}\hat{\phi}\, e^{-\int \mathscr{L}_0}} = \frac{1}{k^2}(2\pi)^D \delta^{(D)}(k+p)\Theta(k) \tag{509}$$

where we define $\Theta(k)$ via

$$\Theta(k) = \begin{cases} 1\,, & b\Lambda \leq |k| < \Lambda \\ 0\,, & \text{otherwise} \end{cases} \tag{510}$$

Peskin and Schroeder denote this correction to the ϕ propagator by 

Let us use what we just learned to help expand the exponential in the $\hat{\phi}$. Terms appear like

$$-\int d^D x \frac{\lambda}{4}\phi^2 \,\overbrace{\hat{\phi}\hat{\phi}}\, = -\frac{1}{2}\int \frac{d^D k}{(2\pi)^D}\,\mu_2\,\phi(k)\phi(-k) \tag{511}$$

where

$$\mu_2 = \frac{\lambda}{2}\int_{b\Lambda\leq|k|<\Lambda} \frac{d^D k}{(2\pi)^D}\frac{1}{k^2} \tag{512}$$

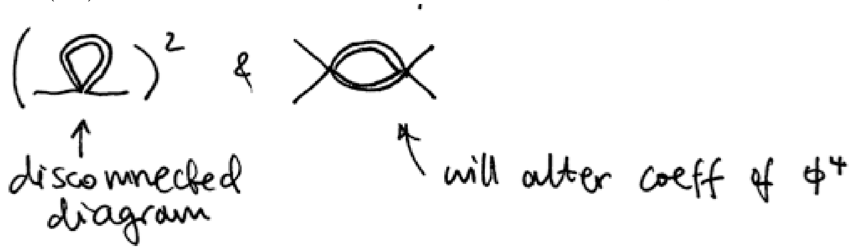
Notice that for $D > 2$ this blows up at large Λ as a power law. Using the definition of a Gamma function and recalling the area formula for spheres in higher dimensions gives

$$\mu_2 = \frac{\lambda\,\Lambda^{D-2}}{(4\pi)^{D/2}\Gamma(\frac{D}{2})}\frac{(1-b^{D-2})}{(D-2)} \tag{513}$$

This would have been the kind of term we were expecting anyway, upon expansion of the exponential in the Feynman path integration over high-energy modes. So μ_2 *gives a positive correction to the mass term in* $\mathscr{L}(\phi)$.

What we have seen here is a great deal more general than the explicit example which we just worked out. Indeed, *integrating out higher-momentum modes shifts around the coupling constants of the low-energy effective action* $\mathscr{L}_{\text{eff}}(\phi)$*!* In particular, changing the UV cutoff makes couplings run and generates nonzero coefficients for all terms allowed by symmetry in the low-energy effective action whether they were present classically or not. This is the modern interpretation of renormalization à la Wilson. It is a hugely powerful idea, and forms the basis for our modern understanding of theoretical high-energy physics and theoretical condensed matter physics as well.

At $\mathscr{O}(\lambda^2)$ there are two possible contractions arising from path integration over $\hat{\phi}$ modes:



The second term contributes to alteration of the coupling constant. In the limit of small external momenta, i.e. for $|k_i^\mu| \ll \Lambda\ \forall i, \mu$, the second diagram is

$$-\frac{1}{4!}\int d^D x\,\zeta\,\phi^4 \tag{514}$$

where

$$\zeta = -4!\frac{2}{2!}\left(\frac{\lambda}{4}\right)^2 \int_{b\Lambda\leq|k|<\Lambda} \frac{d^D k}{(2\pi)^D}\left(\frac{1}{k^2}\right)^2 = -\frac{3\lambda^2\,\Lambda^{D-4}}{(4\pi)^{D/2}\Gamma(\frac{D}{2})}\frac{(1-b^{D-4})}{(D-4)} \tag{515}$$

In the limit that $D \to 4$ this becomes

$$\zeta \to -\frac{3\lambda^2}{16\pi^2}\ln(\frac{1}{b}) \tag{516}$$

If we were to put external momenta back into the equation, rather than neglecting them compared to the UV cutoff, we would generate additional terms of the form

$$-\frac{1}{4}\int d^D x\, \eta\, \phi^2\, (\partial\phi)^2 \tag{517}$$

and higher powers in the Taylor expansion in external momenta.

The message is that *integrating out the $\hat{\phi}$ produces (generates) all possible interactions of fields ϕ and their derivatives consistent with symmetries of the QFT*. Then

$$\mathscr{L}_{\text{eff}}[\phi] = \tfrac{1}{2}(\partial\phi)^2 + \tfrac{1}{2}m^2\phi^2 + \frac{\lambda}{4!}\phi^4 + (\text{connected diagrams}) \tag{518}$$

There is more. We can figure out using the same kind of procedure how changing Λ changes the couplings of the theory. This is known as the RG (renormalization group).

So far, we have seen that at energies well below the UV cutoff our theory is defined via

$$Z = \int [\mathscr{D}\phi]_{b\Lambda} \exp\left(-\int d^D x \mathscr{L}_{\text{eff}}\right) \tag{519}$$

where

$$\begin{aligned} S_{\text{eff}} &= \int d^D x \left\{\frac{1}{2}(1+\Delta Z)(\partial\phi)^2 + \frac{1}{2}(m^2+\Delta m^2)\phi^2 + \right. \\ &\qquad \left. + \frac{1}{4}(\lambda+\Delta\lambda)\phi^4 + \Delta C(\partial\phi)^4 + D\phi^6 + \cdots\right\} \end{aligned} \tag{520}$$

Let us now get rid of the annoying factors of b via a field redefinition. First, we write

$$k' := \frac{1}{b}k\,, \quad x' := bx \tag{521}$$

Then we have that $|k'| \in [0, \Lambda)$. Further,

$$\begin{aligned} S_{\text{eff}} &= \int d^D x' \frac{1}{b^D}\left\{\frac{1}{2}(1+\Delta Z)(\partial'\phi)^2\, b^2 + \frac{1}{2}(m^2+\Delta m^2)\phi^2 + \right. \\ &\qquad \left. + \frac{1}{4}(\lambda+\Delta\lambda)\phi^4 + \Delta C(\partial'\phi)^4\, b^4 + D\phi^6 + \cdots\right\} \end{aligned} \tag{522}$$

This permits us to define

$$\phi' := \sqrt{b^{2-D}(1+\Delta Z)}\phi \tag{523}$$

Then

$$S_{\text{eff}} = \int d^D x' \left\{\frac{1}{2}(\partial'\phi')^2 + +\frac{1}{2}(m')^2(\phi')^2 + \frac{1}{4}\lambda'(\phi')^4 + C'(\partial'\phi')^4 + D'(\phi')^6 + \cdots\right\} \tag{524}$$

So our new parameters in our low-energy effective Lagrangian $\mathscr{L}_{\rm eff}$ are:

$$\begin{aligned}(m')^2 &= (m^2+\Delta m^2)\frac{1}{(1+\Delta Z)}b^{-2} \\ (\lambda') &= (\lambda+\Delta\lambda)\frac{1}{(1+\Delta Z)^2}b^{D-4} \\ C' &= (C+\Delta C)\frac{1}{(1+\Delta Z)^2}b^{D} \\ D' &= (D+\Delta D)\frac{1}{(1+\Delta Z)^3}b^{2D-6}\end{aligned} \tag{525}$$

Let us summarize this extremely important idea:

Cranking down the UV cutoff scale Λ	$\Rightarrow$	Integrating out further high-momentum modes	$\Rightarrow$	FLOW in the space of Lagrangians!!

Note: the RG describing the flow of couplings at the quantum level is mathematically a semigroup, not a group. The reason is that although two lowerings of Λ can be sensibly composed, the transformations do not satisfy all of the group axioms. For this reason, "renormalization group" is a misnomer, and I prefer to call it just the "RG". For a more in-depth discussion, see e.g. Section 9.4 of the text by Tom Banks.

How should we interpret a fixed point? This is simply a point in (multi-dimensional) coupling space at which the RG transformations do not alter the coupling(s). For example, in ϕ^4 field theory, one (trivial) fixed point has

$$\mathscr{L}_0 = \tfrac{1}{2}(\partial\phi)^2\,, \quad m^2=0\,, \quad \lambda=0\,, \quad C=0\,, \quad D=0\,. \tag{526}$$

It is generally useful to study the QFT in the vicinity of the fixed point, to see what kind of flow occurs there. In the vicinity of $\mathscr{L}_0$, Peskin and Schroeder argue that it is valid to ignore the $\Delta m^2, \Delta\lambda$ pieces coming from loop corrections and just focus on the lowest nontrivial order in small quantities. Then

$$(m')^2 = \frac{m^2}{b^2}\,, \quad \lambda' = \lambda b^{D-4}\,, \quad C' = Cb^D\,, \quad D' = Db^{2D-6}\,, \quad \text{etc} \tag{527}$$

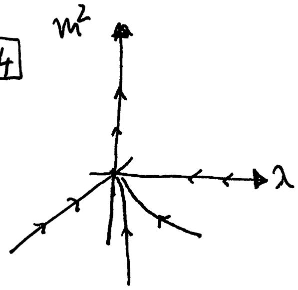
How about power counting in this context?

- Any term in $\mathscr{L}_{\rm eff}$ that comes with a *negative* power of b will *grow* as b is reduced, i.e. at low-energy. These are called *relevant operators.*
- Any term in $\mathscr{L}_{\rm eff}$ that comes with *no* power of b remains unmolested. These are called *marginal[ly relevant] operators.*
- Any term in $\mathscr{L}_{\rm eff}$ that comes with a *positive* power of b will *shrink* at low-energy. These are called (surprise!) *irrelevant operators.*

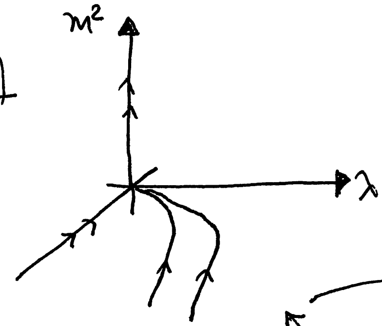
There exists in $D<4$ a fixed point for quartic scalar QFT at nonzero λ known as a Wilson-Fisher fixed point. It exists only for $D<4$ and does not apply in either $D=4$ or $D>4$. This is intimately connected to the observation we made earlier that ϕ^4 theory in $D=4$ is trivial.

Pictures from PS §12.1

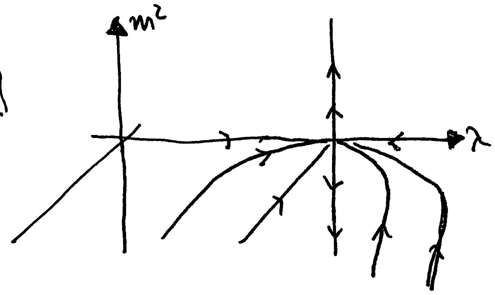
d > 4



d = 4



d < 4



For $d=4$ especially

$$\lambda' = \lambda - \frac{3\lambda^2}{16\pi^2}\log\left(\frac{1}{b}\right) \quad (4.1)$$

which shows λ decreases logarithmically in IR.

Logs typically show up for MARGINAL operators.

(★) For a gorgeous detailed description of these 3 cases, see PS §12.1. ☺ !!

10 One loop renormalization of QED

10.1 Power counting

By dimensional analysis, the superficial degree of divergence $\mathscr{D}$ of a Feynman graph for QED must be

$$\mathscr{D} = DL - 2P_i - E_i \tag{528}$$

where D is the spacetime dimension, L is the number of loops, P_i is the number of internal photon lines and E_i is the number of internal electron lines. The relative factor of 2 between photon and electron internal lines arises because photon propagators scale as $1/k^2$ while fermion propagators scale as $1/k$.

The number of independent loop momenta for integration is

$$L = (\#\text{ internal lines }) - n + 1 \tag{529}$$

because momentum must be conserved at each vertex $(-n)$ and also overall momentum must be conserved as well $(+1)$. Now, how many internal lines are there? Clearly, this number must be the sum of the number of internal photon lines and the number of internal electron lines, because there are no other quantum fields in the theory:

$$L = P_i + E_i - n + 1 \tag{530}$$

Next, let n be the number of vertices, P_e be the number of external photon lines, and E_e be the number of external electron lines. Consider any given vertex in a loop diagram for QED. Each vertex has exactly two electron legs and exactly one photon leg. If an electron leg is external, it is counted once, or, if it is internal, it is counted twice. (This is a simple consequence of the topology of Feynman diagrams. If you are in any doubt, draw an arbitrary Feynman graph in this theory and count its internal and external legs.) Therefore

$$2n = E_e + E_i \tag{531}$$

For photons the analogue is

$$n = P_e + P_i \tag{532}$$

Eliminating internal loop quantities in favour of external ones gives

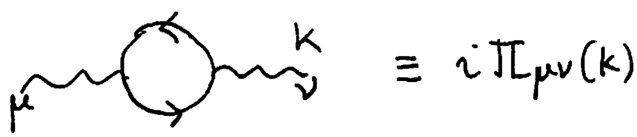
$$\mathscr{D} = D + \left(\frac{D}{2} - 2\right) n - \left(\frac{D-1}{2}\right) E_e - \left(\frac{D-2}{2}\right) P_e \tag{533}$$

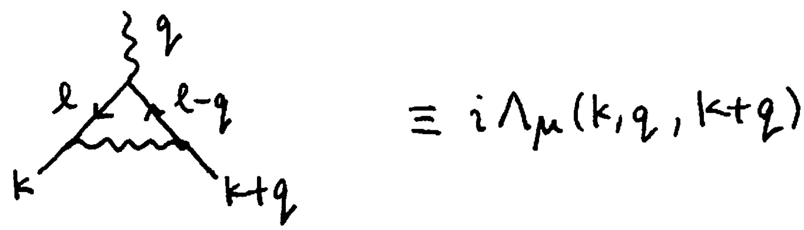
For any given theory to be renormalizable, we need for $\mathscr{D}$ to be independent of n, otherwise we would keep generating new counterterms at each order in perturbation theory. When $D = 4$ we get for QED

$$\mathscr{D}^{D=4}_{\text{QED}} = 4 - \frac{3}{2} E_e - P_e \tag{534}$$

which is, happily, n-independent. However, in $D < 4$ QED is super-renormalizable and strongly coupled in the IR. By contrast, in $D > 4$, QED (as well as generic gauge field theory, in fact) is non-renormalizable and ill-behaved in the UV.

Our next task in attempting to regularize UV infinities in QED and renormalize the theory, we have to chase down where the divergences are. We will focus only on one-loop renormalization in this course, which simplifies our task. If we sit down and write out all possible Feynman diagrams that could be divergent in $D = 4$ QED, we will find a handful. But QED possesses symmetries including CPT, and this ends up reducing the number of possible one-loop divergences in the theory. The result of these symmetry considerations is that there are *only three primitively divergent* diagrams at one loop:





For the photon energy *a.k.a.* vacuum polarization diagram,

$$i\Pi_{\mu\nu}(k) = (-ie)^2 \int \frac{d^D x}{(2\pi)^D} \mathrm{Tr}\gamma^\mu \frac{i}{(\not{l} - m)} \gamma^\nu \frac{i}{(\not{l} + \not{k} - m)} \tag{535}$$

This fellow has $E_e = 0$ and $P_e = 2$ so $\mathscr{D} = 2$ i.e. it is superficially quadratically divergent.

For the electron self-energy, in Feynman gauge we have

$$-i\Sigma(k) = (-ie)^2 \int \frac{d^D x}{(2\pi)^D} \gamma^\mu \frac{i}{(\not{l} - \not{k} - m)} \frac{-i\eta_{\mu\nu}}{l^2} \gamma^\nu \tag{536}$$

This graph has $P_e = 0$ and $E_e = 2$ so $\mathscr{D} = 1$. In other words, this one is superficially linearly divergent.

Both the photon self-energy and the electron self-energy graphs will actually turn out to be *logarithmically* divergent.

The vertex graph at one loop is

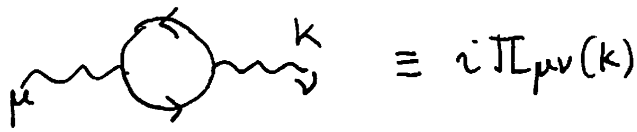
$$-ie\Lambda_\mu(k, q, k+q) = (-ie)^3 \int \frac{d^D x}{(2\pi)^D} \frac{-i\eta_{\rho\sigma}}{(l+k)^2} \gamma^\rho \frac{i}{(\not{l} - \not{q} - m)} \gamma_\mu \frac{i}{(\not{l} - m)} \gamma^\sigma \tag{537}$$

For this guy, we have $E_e = 2$ and $P_e =$ so $\mathscr{D} = 0$. This superficial degree of divergence will turn out to be the same as the actual degree of divergence: logarithmic.

Now we switch to the hunt for logarithmic divergences!

10.2 Photon self-energy a.k.a. Vacuum Polarization

Consider the photon self-energy diagram, which is also known as the vacuum polarization.



We have, in dimensional regularization,

$$\begin{aligned}\Pi_{\mu\nu}(k) &= i\mu^{4-D}e^2\int\frac{d^Dl}{(2\pi)^D}\mathrm{Tr}\left[\gamma_\mu\frac{1}{(\not{l}-m)}\gamma_\nu\frac{1}{(\not{l}-\not{k}-m)}\right]\\ &= ie^2\mu^{4-D}\int\frac{d^Dl}{(2\pi)^D}\frac{\mathrm{Tr}\left[\gamma_\mu(\not{l}+m)\gamma_\nu(\not{l}-\not{k}+m)\right]}{(l^2-m^2)[(l-k)^2-m^2]}\end{aligned} \tag{538}$$

We introduce a Feynman parameter z to gather the two propagator denominators together into one denominator factor, and define a shifted loop momentum l' by

$$l' = l - kz\,. \tag{539}$$

Then

$$\Pi_{\mu\nu}(k) = ie^2\mu^{4-D}\int_0^1 dz\int\frac{d^Dl'}{(2\pi)^D}\frac{\mathrm{Tr}\left[\gamma_\mu(\not{l}'+\not{k}z+m)\gamma_\nu(\not{l}'-\not{k}(1-z)+m)\right]}{[(l')^2-m^2+k^2z(1-z)]^2} \tag{540}$$

Since numerator terms odd in l' do not contribute by symmetry, and since the trace of any gamma matrix is zero, the numerator becomes

$$N_\Pi = \{l'^\kappa l'^\lambda - k^\kappa k^\lambda z(1-z)\}\mathrm{Tr}(\gamma_\mu\gamma_\nu\gamma_\kappa\gamma_\lambda) + m^2\mathrm{Tr}(\gamma_\mu\gamma_\nu) \tag{541}$$

The Peskin and Schroeder appendices contain all dimensional regularization formulæ that we might need for one-loop renormalization of QED. The ones we will use the most in the following are

$$\begin{aligned}\gamma^\mu\gamma_\mu = \eta^\mu_\mu &= D\\ \gamma^\mu\gamma^\nu\gamma_\mu &= (2-D)\gamma_\nu\\ \gamma^\mu\gamma^\nu\gamma^\rho\gamma_\mu &= 4\eta^{\nu\rho} + (D-4)\gamma^\nu\gamma^\rho\\ \gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\gamma_\mu &= -2\gamma^\sigma\gamma^\rho\gamma^\nu + (4-D)\gamma^\nu\gamma^\rho\gamma^\sigma\\ \mathrm{Tr}(\gamma^\mu\gamma^\nu) &= 4\eta^{\mu\nu}\\ \mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) &= 4(\eta^{\mu\rho}\eta^{\nu\sigma} - \eta^{\mu\nu}\eta^{\rho\sigma} + \eta^{\mu\lambda}\eta^{\nu\sigma})\end{aligned} \tag{542}$$

The rationale behind these formulæ is that gamma matrix traces not involving γ_5 are unaltered from their $D=4$ versions – except for those traces involving $\eta^\mu_\mu = D$.

We will also find useful the following symmetry relations for tensor numerators in one-loop integrals:

$$l^\mu l^\nu \longrightarrow \frac{1}{D}l^2\eta^{\mu\nu}$$

$$l^\mu l^\nu l^\rho l^\sigma \quad \longrightarrow \quad \frac{1}{D(D+2)}(l^2)^2\left[\eta^{\mu\nu}\eta^{\rho\sigma}+\eta^{\mu\rho}\eta^{\nu\sigma}+\eta^{\mu\sigma}\eta^{\nu\rho}\right] \tag{543}$$

Using the first of these identities, our numerator for the vacuum polarization becomes

$$N_\Pi \;=\; 4[l'^\kappa l'^\lambda - k^\kappa k^\lambda z(1-z)](\eta_{\mu\kappa}\eta_{\nu\lambda}-\eta_{\mu\nu}\eta_{\kappa\lambda}+\eta_{\nu\kappa}\eta_{\mu\lambda})+4m^2\eta_{\mu\nu} \tag{544}$$

Therefore

$$\begin{aligned}\frac{N_\Pi}{4} &= \eta_{\mu\nu}\left[m^2-(l')^2+k^2z(1-z)\right]+l'_\mu l'_\nu+l'_\nu l'_\mu-(k_\mu k_\nu+k_\mu k_\nu)z(1-z)\\ &= 2l'_\mu l'_\nu-2\left[k_\mu k_\nu-\eta_{\mu\nu}k^2\right]z(1-z)-\eta_{\mu\nu}\left[(l')^2-m^2+k^2z(1-z)\right]\end{aligned} \tag{545}$$

Here we have added and subtracted a term for clarity's sake. Using also the symmetry identity $l^\mu l^\nu \longrightarrow (1/D)\eta^{\mu\nu}l^2$, we see that the first and third terms cancel in the integral:

$$\begin{aligned}\Pi_{\mu\nu}(k) &= 4ie^2\mu^{4-D}\int_0^1 dz\int\frac{d^Dx}{(2\pi)^D}\left\{\frac{2l'_\mu l'_\nu}{[l^2-m^2+k^2z(1-z)]^2}\right.\\ &\left.-\frac{2z(1-z)}{[l^2-m^2+k^2z(1-z)]^2}-\frac{\eta_{\mu\nu}}{[l^2-m^2+k^2z(1-z)]^1}\right\}\end{aligned} \tag{546}$$

Therefore,

$$\Pi_{\mu\nu}(k) = -8ie^2\mu^{4-D}\int_0^1 dz\int\frac{d^Dl}{(2\pi)^D}\frac{z(1-z)(k_\mu k_\nu-\eta_{\mu\nu}k^2)}{[l^2-m^2+k^2z(1-z)]} \tag{547}$$

We now use a formula from PS (A.44) for the kind of loop momentum integral we are doing to compute the one-loop photon self-energy:

$$\int\frac{d^Dl}{(2\pi)^D}\frac{1}{(l^2-\Delta)^n} = \frac{(-1)^n i}{(4\pi)^{D/2}}\frac{\Gamma(n-D/2)}{\Gamma(n)}\frac{1}{\Delta^{n-D/2}} \tag{548}$$

In order to take the $\varepsilon = (4-D)\to 0$ limit, it is useful to know another formula from PS appendix (A.49):

$$\lim_{D\to 4}\frac{1}{\Delta^{2-D/2}} = 1+(\frac{D}{2}-2)\log\Delta+\dots \tag{549}$$

Peskin and Schroeder also state a formula for a particular combination of factors which often turns out to be relevant to interesting one-loop QFT calculations: (A.52). This little treasure allows us to correctly get the 4π factor in the following:

$$\lim_{D\to 4}\frac{\Gamma(2-D/2)}{(4\pi)^{D/2}}\frac{1}{\Delta^{2-D/2}} = \frac{1}{(4\pi)^2}\left[\frac{2}{\varepsilon}-\log\Delta-\gamma+\log(4\pi)+\mathscr{O}(\varepsilon)\right] \tag{550}$$

The other useful formula which we will use multiple times is PS equation (A.51) for how to expand a Gamma function near its poles:

$$\Gamma(x) = \frac{(-1)^n}{n!}\left(\frac{1}{x+n}-\gamma+1+\dots+\frac{1}{n}+\mathscr{O}(x+n)\right) \tag{551}$$

where γ is the Euler-Mascharoni constant.

Putting all these ingredients together, we obtain finally

$$\Pi_{\mu\nu}^{D=4}(k) = \frac{e^2}{2\pi^2}(k_\mu k_\nu - \eta_{\mu\nu}k^2)\left\{\frac{1}{3\varepsilon} - \frac{\gamma}{6} - \int_0^1 dz \log\left(\frac{k^2 z(1-z) - m^2}{4\pi\mu^2}\right) + \text{finite}\right\} \tag{552}$$

Note that this is proportional to the original tree-level photon propagator – we will use this neat fact when we talk about charge renormalization.

One final remark about $\Pi_{\mu\nu}(k)$. Notice that the maximum value of $z(1-z)$ on the interval $z \in [0,1]$ is $1/4$, at $z = 1/2$. Therefore, we can see that the log function will develop a branch cut whenever $k^2 > 4m^2$. Because the amount of energy needed to create a pair is $2m$ in units where $c = 1$, development of a branch cut actually signals pair production. This is a nice physical consistency check on our complicated mathematical formulæ.

10.3 Electron self-energy

Consider the one loop correction to the electron propagator in QED

It takes the form

$$\frac{i(\not{k} + m)}{k^2 - m^2}\left[-i\Sigma_2(k)\right]\frac{i(\not{k} + m)}{k^2 - m^2} \tag{553}$$

where the amputated part of the vertex $-i\Sigma_2(k)$ is given by

$$-i\Sigma_2(k) = (-ie)^2 \int \frac{d^D l}{(2\pi)^D}\, \gamma^\mu\, \frac{i(\not{l} + m)}{l^2 - m^2 + i\epsilon}\, \gamma_\nu\, \frac{-i\eta^{\mu\nu}}{[(l-k)^2 + i\epsilon]} \tag{554}$$

As we did with the photon self-energy, our next step is to use the Feynman parameter technique to collect the propagator denominators into one combined denominator so we can do the integral. We use

$$\frac{1}{(l^2 - m^2 + i\epsilon)[(k-l)^2 + i\epsilon]} = \int_0^1 dz \frac{1}{[l^2 - 2zl\cdot k + zk^2 - (1-z)m^2 + i\epsilon]} \tag{555}$$

Next, we complete the square and define a shifted loop momentum

$$l' = l - zk \tag{556}$$

Using these two identities, and doing a Wick rotation to Euclidean loop momenta, we obtain

$$-i\Sigma_2(k) = -e^2 \int_0^1 dz \int \frac{d^D l_E}{(2\pi)^D} \frac{[-2(2-\varepsilon)\not{k} + (4-\varepsilon)m]}{\left[l_E^2 + \Delta - i\varepsilon\right]^2} \tag{557}$$

where we dropped the primes on the loop momentum, used symmetry to drop numerator terms linear in l, and defined

$$\Delta = -z(1-z)k^2 + (1-z)m^2\,. \tag{558}$$

But in dimensional regularization we have

$$\int \frac{d^D l_E}{(2\pi)^D} \frac{1}{(l_E^2 + \Delta - i\varepsilon)^2} = \frac{1}{(4\pi)^{D/2}} \frac{\Gamma(2 - D/2)}{\Delta^{2-D/2)}} \tag{559}$$

Also,

$$\lim_{D\to 4} \frac{\Gamma(2 - D/2)}{(4\pi)^{D/2}} \frac{1}{\Delta^{2-D/2}} = \frac{1}{(4\pi)^2}\left[\frac{2}{\varepsilon} - \log\Delta - \gamma + \log(4\pi) + \mathscr{O}(\varepsilon)\right] \tag{560}$$

Putting the components together, we have

$$\begin{aligned}
\Sigma_2(k) &= \frac{e^2}{(4\pi)^{D/2}} \frac{\Gamma(2 - D/2)}{\Delta^{2-D/2}} \left\{(4-\varepsilon)m - \frac{1}{2}(2-\varepsilon)\not{k}\right\} \\
&= \frac{e^2}{(4\pi)^{D/2}} \frac{\Gamma(2 - D/2)}{\Delta^{2-D/2}} \left\{m(3 - \frac{\varepsilon}{2}) + (\not{k} - m)(\frac{\varepsilon}{2} - 1)\right\} \\
&\to \frac{e^2}{16\pi^2}\left[\frac{2}{\varepsilon} - \log(\frac{\Delta}{4\pi\mu^2}) - \gamma + \mathscr{O}(\varepsilon)\right]\left\{(4m - \not{k}) + \varepsilon(\frac{\not{k}}{2} - m)\right\} \\
&= \frac{e^2}{8\pi^2}(4m - \not{k})\frac{1}{\varepsilon} + \text{finite}
\end{aligned} \tag{561}$$

Notice how factors involving $1/\varepsilon$ and ε cancel in the finite part of this expression. This is one reason why we have been very careful to keep ε-dependent pieces in the numerator!

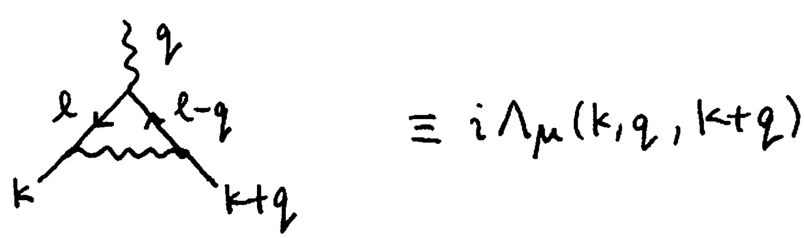
Therefore, the electron self-energy at 1-loop modifies the electron's inverse propagator to

$$\begin{aligned}
\Gamma^{(2)}(p) &= S_F'(p)^{-1} = S_F(p)^{-1} - \Sigma(p) \\
&= (\not{p} - m) - \frac{e^2}{8\pi^2\varepsilon}(-\not{p} + 4m) + \text{finite} \\
&= \not{p}\left(1 + \frac{e^2}{8\pi^2\varepsilon}\right) - m\left(1 + \frac{e^2}{2\pi^2\varepsilon}\right) + \text{finite}
\end{aligned} \tag{562}$$

10.4 QED Vertex Correction

First, a parenthetical note. There are actually other Feynman diagrams to worry about in QED than the UV-divergent loop contributions from self-energies or vertex corrections. For a *long-range* theory like electromagnetism there are also IR-divergent diagrams, which essentially correspond to radiating soft photons off external leg lines. We do not discuss IR divergences here, but encourage serious students of QFT to learn what a "Sudakov log" is.

For the vertex correction in QED we are concerned with this diagram:



Getting back to our UV-divergent one-loop vertex, we have

$$i\Lambda_\mu(k, q, k+q) = -ie^2 \int \frac{d^D l}{(2\pi)^D} \frac{\gamma_\rho\,(\not{l} + m)\,\gamma^\mu\,(\not{l} - \not{q} + m)\,\gamma^\rho}{[(l-k)^2 + i\epsilon]\,[l^2 - m^2 + i\epsilon]\,[(l-q)^2 - m^2 + i\epsilon]} \tag{563}$$

Oh dear. We are slightly stuck here if we think in terms of techniques we used for the two self-energy diagrams, because of the third denominator. So how do we collect three or more denominators together? An extremely useful general formula, which can be derived by induction as in PS p.190 (section 6.3), is:

$$\frac{1}{A_1^{m_1} A_2^{m_2} \cdots A_n^{m_n}} = \int_0^1 dz_1 \int_0^1 dz_2 \cdots \int_0^1 dz_n \; \delta(\sum_i z_i - 1) \times \times \frac{\prod_i z_i^{m_i - 1}}{[\sum_j z_j A_j]^{\sum_k m_k}} \frac{\Gamma(m_1 + m_2 + \cdots + m_n)}{\Gamma(m_1)\Gamma(m_2)\cdots\Gamma(m_n)} \tag{564}$$

For the three-denominator case, we need

$$\frac{1}{[(l-k)^2 + i\epsilon]\,[l^2 - m^2 + i\epsilon]\,[(l-q)^2 - m^2 + i\epsilon]} = \int_0^1 dxdydz \; \delta(x+y+z-1)\, \frac{2}{D^3} \tag{565}$$

where

$$D = x[(l-k)^2] + y[l^2 - m^2] + z[(l-q)^2 - m^2] + (x+y+z)i\epsilon \tag{566}$$

By changing to a new dummy variable of loop integration

$$l' = (l-q) + yq - x(k-q) \tag{567}$$

and using $x + y + z = 1$ it is possible – though boring – to show that the denominator collapses to

$$D = (l')^2 - \Delta + i\epsilon \tag{568}$$

where

$$\Delta = -xyq^2 + (1-z)^2 m^2 \tag{569}$$

Notice that for physical scattering processes in Lorentzian signature $q^2 < 0$ so no branch cuts 'plague' us here.

Then, dropping the primes on l', we have

$$\Lambda_\mu(q,k,k-q) = ie^2 \int_0^1 dx \int_0^{1-x} dy \int \frac{d^D l}{(2\pi)^D} \, \gamma_\nu \, (\not{l} - m) \, \gamma^\mu \, [(\not{l} - \not{q} + y\not{q} - x(\not{k} - \not{q}) - m]\gamma^\nu \frac{2}{D^3} \tag{570}$$

The next step is to use symmetry to kill any terms odd in the loop momentum. We also use the the dimensional regularization Dirac identities to cope with all the gamma matrices. The part of the numerator quadratic in l is divergent, and has the value

$$\Lambda_\mu(q,k,k-q) = \frac{e^2}{8\pi^2}\frac{1}{\varepsilon}\gamma_\mu + \text{finite} \tag{571}$$

The finite piece of this ends up contributing to the anomalous magnetic moment of the electron, but we will not chase that fact up here.

(Note: the definition of Γ_μ we use is such that it has an implicit factor of e, from the tree vertex, inside. That way the one-loop correction is proportional to e^2, not e^3.)

10.5 Ward-Takahashi Identities in QED

We have learned so far that for connected graphs we use $W[J]$, defined by

$$Z[J] = \exp\left(iW[J]\right) \tag{572}$$

where $Z[J]$ is the original generating functional. We also learned that for one particle irreducible (1PI) graphs we focus on the quantum action, and for the case of the scalar $\Gamma[\phi]$ was given by

$$W[J] = \Gamma[\phi] + \int d^D x J(x)\phi(x) \tag{573}$$

Consider now QED with its $U(1)$ gauge symmetry. Without the gauge-fixing and ghost terms, our Lagrangian was gauge invariant. Adding $\mathscr{L}_{\text{gf}} + \mathscr{L}_{\text{gh}}$ to the story tamed the measure of the Feynman Path Integral but made it non-gauge-invariant!! So does Z have some secret gauge invariance? Certainly, scattering amplitudes should not depend on the gauge chosen. Therefore, we can conclude that Z must obey some relations. These will end up being the Ward-Takahashi Identities.

We have

$$Z = n \int \mathscr{D}A_\mu \mathscr{D}\bar{\psi}\mathscr{D}\psi \exp\left(i \int \mathscr{L}_{\text{eff}}\right) \tag{574}$$

where our effective Lagrangian includes the mandatory source terms:

$$\mathscr{L}_{\text{eff}} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + i\bar{\psi}\gamma^\mu\left(\partial_\mu + ieA_\mu\right)\psi - m\bar{\psi}\psi - \frac{1}{2\alpha}\left(\partial^\mu A_\mu\right)^2 + J^\mu A_\mu + \bar{\eta}\psi + \bar{\psi}\eta \tag{575}$$

Suppose that we do a(n infinitesimal) gauge transformation

$$\begin{aligned} A_\mu &\to A_\mu + \partial_\mu\lambda \\ \psi &\to \psi - ie\lambda\psi \\ \bar{\psi} &\to \bar{\psi} + ie\lambda\bar{\psi} \end{aligned} \tag{576}$$

By construction, the first three terms are gauge invariant. The remaining pieces are the prime suspects: the gauge-fixing Lagrangian, the ghost Lagrangian, and the source terms.

Under (finite) gauge transformations, the FPI integrand will pick up the factor

$$\exp\left(i\int d^D x\left[-\frac{1}{\alpha}(\partial^\mu A_\mu)\partial^2\lambda - \partial^\mu J_\mu - ie\lambda(\bar{\eta}\psi - \bar{\psi}\eta)\right]\right) \tag{577}$$

For infinitesimal $\lambda(x)$, this expands as

$$\left\{1 + i\int d^D x\left[-\frac{1}{\alpha}\partial^2(\partial^\mu A_\mu) - \partial^\mu J_\mu - ie(\bar{\eta}\psi - \bar{\psi}\eta)\right]\lambda(x)\right\} \tag{578}$$

after some integrations by parts.

Invariance of the full Z under these transformations would require, infinitesimally,

$$\left\{1 + i\int d^D x\left[-\frac{1}{\alpha}\partial^2(\partial^\mu A_\mu) - \partial^\mu J_\mu - ie(\bar{\eta}\psi - \bar{\psi}\eta)\right]\lambda(x)\right\} Z = Z \tag{579}$$

for arbitrary $\lambda(x)$. Therefore,

$$\left[-\frac{1}{\alpha}\partial^2(\partial\cdot A) - \partial\cdot J - ie(\bar{\eta}\psi - \bar{\psi}\eta)\right] Z = 0 \tag{580}$$

Next, by the source dependence of $Z[J,\eta,\bar{\eta}]$, we can substitute

$$\begin{aligned} \psi &\rightarrow \frac{1}{i}\frac{\delta}{\delta\bar{\eta}} \\ \bar{\psi} &\rightarrow \frac{1}{i}\frac{\delta}{\delta\eta} \\ A_\mu &\rightarrow \frac{1}{i}\frac{\delta}{\delta J^\mu} \end{aligned} \tag{581}$$

This gives a *functional differential equation*

$$\left[\frac{i}{\alpha}\partial^2\partial^\mu\frac{\delta}{\delta J^\mu} - \partial^\mu J_\mu - e\left(\bar{\eta}\frac{\delta}{\delta\bar{\eta}} - \eta\frac{\delta}{\delta\eta}\right)\right] Z[J,\eta,\bar{\eta}] = 0 \tag{582}$$

For physical Feynman graph applications it is typically more useful to use the connected generating functional, so we substitute

$$Z[J,\eta,\bar{\eta}] = \exp\left(iW[J,\eta,\bar{\eta}]\right) \tag{583}$$

which gives for W

$$\frac{1}{\alpha}\partial^2\partial^\mu\frac{\delta W}{\delta J^\mu} - \partial^\mu J_\mu - ie\left(\bar{\eta}\frac{\delta W}{\delta\bar{\eta}} - \eta\frac{\delta W}{\delta\eta}\right) = 0 \tag{584}$$

The final step is to do our field Legendre transformation to get to the quantum action Γ. Our generating functional for vertex functions $\Gamma^{(n)}$ is defined by

$$\Gamma[A_\mu,\psi,\bar{\psi}] = W[J,\eta,\bar{\eta}] - \int d^Dx\left(J^\mu A_\mu + \bar{\eta}\psi + \bar{\psi}\eta\right) \tag{585}$$

and therefore

$$\begin{aligned} \frac{\delta\Gamma}{\delta A_\mu(x)} &= -J(x) & \frac{\delta W}{\delta J_\mu(x)} &= A_\mu(x) \\ \frac{\delta\Gamma}{\delta\psi(x)} &= -\bar{\eta}(x) & \frac{\delta W}{\delta\bar{\eta}(x)} &= \psi(x) \\ \frac{\delta\Gamma}{\delta\bar{\psi}(x)} &= -\eta(x) & \frac{\delta W}{\delta\eta(x)} &= \bar{\psi}(x) \end{aligned} \tag{586}$$

Accordingly, our overall demand of gauge invariance for the FPI gives the condition

$$-\frac{1}{\alpha}\partial^2\partial^\mu A_\mu(x) + \partial_\mu\frac{\delta\Gamma}{\delta A_\mu(x)} - ie\psi\frac{\delta\Gamma}{\delta\psi(x)} + ie\bar{\psi}\frac{\delta\Gamma}{\delta\bar{\psi}(x)} = 0 \tag{587}$$

This is nice, but it doesn't ring any bells physically! To beat this expression into something useful, we can try an old standard: [functionally] differentiating it. Specifically, we will apply

$$\frac{\delta}{\delta\bar{\psi}(x_1)}\frac{\delta}{\delta\psi(y_1)} \tag{588}$$

and then set the sources to zero and see what eventuates. Note that because $\mathscr{L}_{\rm gf}$ does not depend on $\psi, \bar{\psi}$, the first term will vanish. Our Ward-Takahashi Identity then takes the form

$$
\begin{aligned}
-\frac{\partial}{\partial x^\mu}\left\{\frac{\delta^3[0]}{\delta\bar{\psi}(x_1)\delta\psi(y_1)}\delta A^\mu(x)\right\} \quad = \quad & +ie\delta^D(x-x_1)\frac{\delta^2\Gamma[0]}{\delta\bar{\psi}(x_1)\delta\psi(y_1)} \\
& -ie\delta^D(x-y_1)\frac{\delta^2\Gamma[0]}{\delta\bar{\psi}(x_1)\delta\psi(y_1)}
\end{aligned}
\tag{589}
$$

The LHS of this expression is the derivative of the 1PI electron-positron-photon vertex, while the RHS terms are the inverses of exact propagators!!

The physical content of our WTI equation is a lot easier to digest in momentum (Fourier) space. So let us define

$$
\begin{aligned}
&\int d^Dx \int d^Dx_1 \int d^Dy_1 e^{i(p'\cdot x_1 - p\cdot y_1 - qx)}\frac{\delta^3\Gamma[0]}{\delta\bar{\psi}(x_1)\delta\psi(y_1)\delta A^\mu(x)} \\
&:= ie(2\pi)^D\delta^D(p'-p-q)\Gamma_\mu(p,p',q)
\end{aligned}
\tag{590}
$$

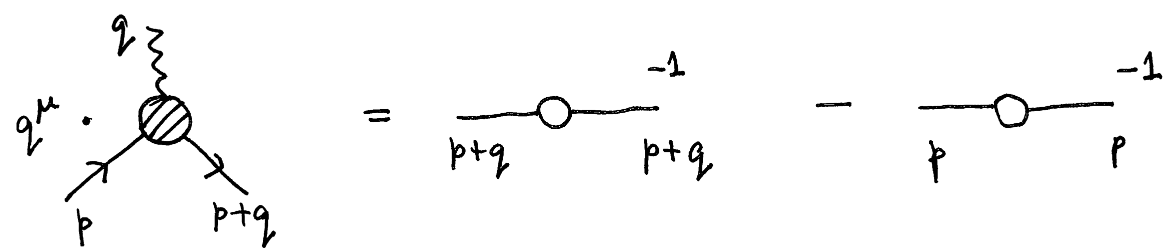
Let us also define the exact propagator – as distinct from the bare one – by

$$
\int d^Dx_1 \int d^Dy_1 e^{i(p'\cdot x_1 - p\cdot y_1}\frac{\delta^2\Gamma[0]}{\delta\bar{\psi}(x_1)\delta\psi(y_1)} := (2\pi)^D\delta^D(p'-p) \ \times \ i(S_F')^{-1}(p) \tag{591}
$$

Now, multiply our position-space WTI by $e^{i(p'\cdot x_1 - p\cdot y_1 - q\cdot x)}$ and integrate w.r.t. x_1, y_1, x. The resulting WTI is

$$
q^\mu\Gamma_\mu(p, p' = p+q, q) = (S_F')^{-1}(p+q) - (S_F')^{-1}(p) \tag{592}
$$

This momentum space WTI is the most useful version of it. Diagrammatically, we have derived



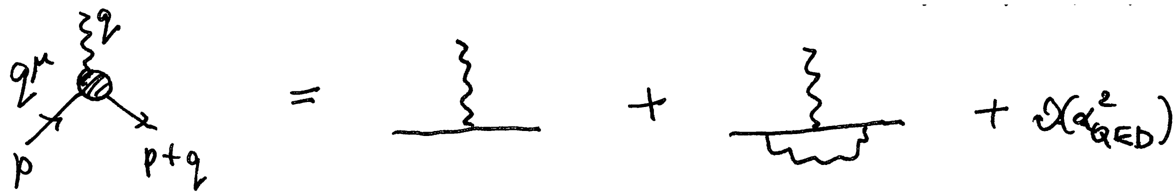
If we take the zero-momentum limit of the momentum space WTI, we obtain

$$
\frac{\partial}{\partial p^\mu}(S_F')^{-1}(p) = \Gamma_\mu(p,p,0) \tag{593}
$$

which is known as the Ward Identity.

Ward-Takahashi identities are *exact* – in other words, true to all orders in perturbation theory. So at any given order in $\alpha_{\rm QED}$ there will be nontrivial relationships between exact propagators and vertex functions!

To see what sort of story eventuates, we can imagine all possible connected, 1PI loop Feynman graphs built from bare propagators and vertices. Diagrammatically,



On the other hand, for the exact propagator for the electron-positron field:

Lastly, let us work out some of the details of the lowest-order WTI. We have the bare fermion propagator

$$S_F^{(0)} = \frac{1}{\not p - m} \tag{594}$$

so its inverse is $(S_F^{(0)})^{-1}(p) = \not p - m$. The momentum derivative is, at leading order,

$$\frac{\partial}{\partial p^\mu}(S_F^{(0)})^{-1}(p)\bigg|_{\text{LO}} = \gamma_\mu \tag{595}$$

The question is: what sits on the other side of the WTI from this term?

We need to find the third derivative of Γ w.r.t. $\psi, \bar\psi, A_\mu$. Using the definitions of Z, W, Γ and the knowledge that at lowest nontrivial order the inverse propagators are $\Gamma^{(2)}$s gives

$$\begin{aligned} \frac{\delta^3\Gamma}{\delta\bar\psi(x_1)\delta\psi(y_1)\delta A^\mu(x)} &= -\int du_1 dv_1 \left\{ iS_F^{-1}(u_1 - x_1) iS_F^{-1}(v_1 - y_1)\right\} \times \\ &\times \left[-iD_{\mu\nu}^{-1}(u-x)(-i)\frac{\delta^3 Z[0]}{\delta\eta(u_1)\delta\bar\eta(v_1)\delta J^\nu(u)}\right] \end{aligned} \tag{596}$$

Our next step is to use the form of $\mathscr{L}_{\text{tot}}$. For QED we have the interaction Lagrangian

$$\mathscr{L}_{\text{int}} = e\bar\psi\gamma^\mu\psi A_\mu \tag{597}$$

and so

$$Z[\eta,\bar\eta,J^\mu] = \mathscr{N}\exp\left(ie\int d^Dz \frac{-i\delta}{\delta\eta(z)}\gamma^\lambda \frac{-i\delta}{\delta\bar\eta(z)}\frac{-i\delta}{\delta J^\lambda(z)}\right) Z_0 \tag{598}$$

where Z_0 is the free version

$$Z_0 = \exp\left(-i\int d^Dx d^Dy \bar\eta(x) S_F(x-y)\eta(y)\right) \times \exp\left(\frac{i}{2}\int d^Dx d^Dy J^\mu(x) D_{\mu\nu}(x-y) J^\nu(y)\right) \tag{599}$$

Therefore, to lowest nontrivial order in α_{QED}, we have

$$\frac{\delta^3 Z[0]}{\delta\eta(u_1)\delta\bar\eta(v_1)\delta A^\mu(u)} = ie\int d^Dz S_F(u_1 - z) S_F(v_1 - z) D_{\mu\nu}(u-z)\gamma^\nu \tag{600}$$

Substituting in to our general form of the WTI of QED gives its lowest-order approximation:

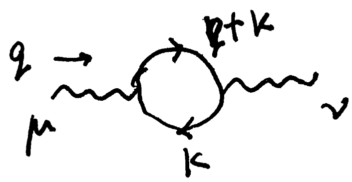
$$\Gamma_\mu(p, p+q, q)\big|_{\rm LO} = \gamma_\mu \tag{601}$$

which is independent of the photon momentum q^μ. Note that this does not generalize to higher loops. We see that the WTI is more than adequately satisfied at lowest order in QED. (For this we did not even need to use the form of $D_{\mu\nu}$.)

10.6 Photon masslessness and Charge Renormalization

Previously, we learned about propagator corrections arising from one-loop contributions. We saw that the massive scalar field of $\lambda\phi^4$ theory suffered a mass shift at one loop. Here, in the QED context, we will see that the photon actually stays massless - *to all orders in perturbation theory.* This is a direct consequence of the Ward Identity which, as we just saw, is itself a consequence of $U(1)$ gauge invariance.

Consider the exact photon propagator. This is physically relevant to any process involving photons. At one loop, we had

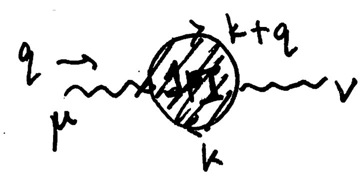


$$\equiv i\Pi_2^{\mu\nu}(q)$$

$$i\Pi_2^{\mu\nu}(q) = (-ie)^2(-1)\int \frac{d^Dk}{(2\pi)^D}\text{Tr}\left[\gamma^\mu \frac{i}{(\not{k}-m)}\gamma^\nu \frac{i}{[(\not{k}+\not{q}-m]}\right] \tag{602}$$

where the factor of (-1) comes from the fermion loop.

For the general physical process we need the *exact, 1PI* propagator contributions



$$\equiv i\Pi^{\mu\nu}(q)$$

The Ward Identity says that

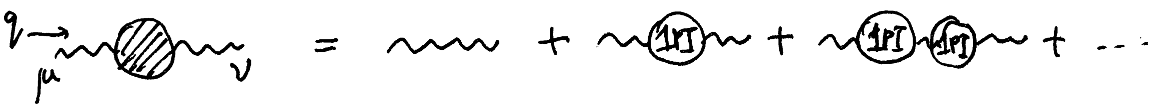
$$q_\mu \Pi^{\mu\nu} = 0 \tag{603}$$

Now, the only tensor structures buildable out of $\eta_{\mu\nu}$ and q_λ are $\eta_{\mu\nu}$ and $q_\mu q_\nu/q^2$. Requiring orthogonality between q and Π implies that

$$\Pi^{\mu\nu}(q) = \left[q^2\eta^{\mu\nu} - q^\mu q^\nu\right]\Pi(q) \tag{604}$$

Peskin and Schroeder argue (p.245) in some detail that $\Pi(q^2)$ should be regular at $q=0$. Essentially, this follows from the structure of 1PI diagrams, which cannot have massless intermediate states appearing in them.

Pick Lorentz-Feynman gauge. The exact photon two-point function is then composed of



$$
\begin{aligned}
i\Pi^{\mu\nu}(q) &= -i\eta_{\mu\nu}\frac{1}{q^2} + \left(-i\eta_{\mu\rho}\frac{1}{q^2}\right)\left[i(q^2\eta^{\rho\sigma} - q^\rho q^\sigma)\Pi(q^2)\right]\left(-i\eta_{\sigma\nu}\frac{1}{q^2}\right) + \ldots \\
&= -i\eta_{\mu\nu}\frac{1}{q^2} + -i\eta_{\mu\rho}\frac{1}{q^2}\Delta^\rho_\nu\Pi(q^2) + -i\eta_{\mu\rho}\frac{1}{q^2}\Delta^\rho_\sigma\Delta^\sigma_\nu\Pi^2(q^2) + \ldots
\end{aligned}
\tag{605}
$$

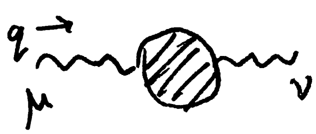
where Δ is defined as

$$
\Delta^\mu_\nu = \delta^\mu_\nu - \frac{q^\mu q_\nu}{q^2}
\tag{606}
$$

As you should check, this tensor Δ is a projector. Therefore,

$$
\begin{aligned}
i\Pi^{\mu\nu}(q^2) &= \left(-i\eta_{\mu\rho}\frac{1}{q^2}\right)\left(\eta^\rho_\nu - \frac{q^\rho q_\nu}{q^2} + \frac{q^\rho q_\nu}{q^2}\right) + \left(-i\eta_{\mu\rho}\frac{1}{q^2}\right)(\Delta^\rho_\nu)\left[\Pi(q^2) + \Pi^2(q^2) + \ldots\right] \\
&= -\frac{i}{q^2}(\Delta_{\mu\nu})\left[1 + \Pi(q^2) + \Pi^2(q^2) + \ldots\right] - \frac{iq_\mu q^\nu}{q^4} \\
&= \frac{-i\Delta_{\mu\nu}}{q^2(1-\Pi(q^2))} - \frac{iq_\mu q^\nu}{q^4} \\
&= \frac{-i\eta_{\mu\nu}}{q^2[1-\Pi(q^2)]} + \frac{iq_\mu q_\nu}{q^4[1-\Pi(q^2)]} - \frac{iq_\mu q^\nu}{q^4} \\
&= \frac{-i\eta_{\mu\nu}}{q^2[1-\Pi(q^2)]} + \frac{iq_\mu q_\nu}{q^4[1-\Pi(q^2)]}\left\{1 - [1-\Pi(q^2)]\right\} \\
&= \frac{-i\eta_{\mu\nu}}{q^2[1-\Pi(q^2)]} + \frac{iq_\mu q_\nu\Pi(q^2)}{q^4[1-\Pi(q^2)]}
\end{aligned}
\tag{607}
$$

Now, in any S-matrix computation, at least one end of our exact photon propagator will connect to a fermion line. Therefore, summing over all places along the line where it could connect, we must find that terms proportional to $q_\mu q_\nu$ vanish by the Ward Identiy. So we will abbreviate the above expression for the exact photon propagator – valid for S-matrix elements only! – as



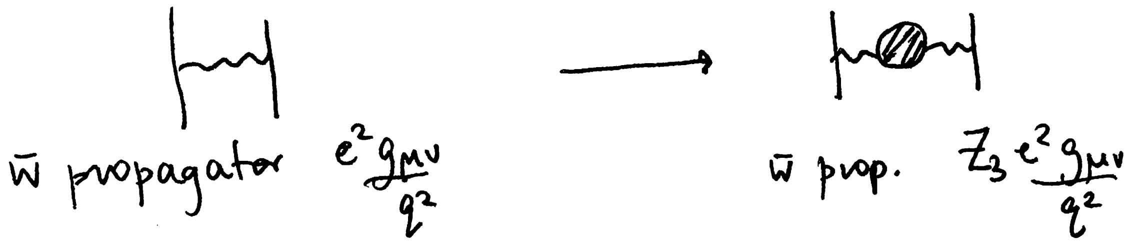
$$
= \frac{-i\eta_{\mu\nu}}{q^2[1-\Pi(q^2)]}
\tag{608}
$$

Look closely at this form of the propagator. Notice that no mass shift appears: there is no loop-generated correction to the tree level result. This masslessness of the photon is a direct consequence of the structure of QED including the Ward Identity, as we have just seen.

The only pole in the exact photon propagator is at $q^2 = 0$. The residue of this pole at $q^2 = 0$ is

$$
\frac{1}{[1-\Pi(0)]} := Z_3
\tag{609}
$$

Any low-q^2 scattering process gets this shift. For example,



At this $\mathscr{O}(\alpha)$ level of approximation, then, a scattering amplitude at *nonzero* q^2 would involve

$$\frac{-i\eta_{\mu\nu}}{q^2}\left(\frac{e_0^2}{1-\Pi(q^2)}\right) \simeq \frac{-i\eta_{\mu\nu}}{q^2}\frac{\alpha}{1-[\Pi_2(q^2)-\Pi_2(0)]} \tag{610}$$

Therefore, at $\mathscr{O}(\alpha)$, we may identify

$$\alpha_{\text{eff}}(q^2) = \frac{e_0^2/(4\pi)}{1-\Pi(q^2)} \simeq \frac{\alpha}{1-[\Pi_2(q^2)-\Pi_2(0)]} \tag{611}$$

Peskin and Schroeder explain later in their textbook that, in the above form, this expression is true to all orders (!) when Π_2 is replaced by the full Π.

We now look back a few subsections to see what we actually found for the one-loop photon propagator correction a.k.a. the vacuum polarization. We had

$$i\Pi_2^{\mu\nu}(q^2) = (q^2\eta^{\mu\nu} - q^\mu q^\nu)i\Pi_2(q^2) \tag{612}$$

where

$$\Pi_2(q^2) = \frac{-2\alpha}{\pi}\int_0^1 dx x(1-x)\left[\frac{2}{(4-d)} - \log\Delta - \gamma + \log(4\pi)\right] \tag{613}$$

The divergent part of this is x-independent, so the Feynman integral for the first term above collapses to 1/6. Accordingly, at $\mathscr{O}(\alpha)$,

$$\Pi_2(0) = \frac{-2\alpha}{3\pi(4-D)} \tag{614}$$

which blows up as $D \to 4$. This is not a worry physically, because this infinite quantity is not actually observed. What is physically measurable is

$$\hat{\Pi}_2(q^2) := \Pi_2(q^2) - \Pi_2(0) = -\frac{2\alpha}{\pi}\int_0^1 dx x(1-x)\log\left(\frac{m^2}{m^2 - x(1-x)q^2}\right) \tag{615}$$

which is totally independent of $(4-D)$ as $D \to 4$.

10.7 The Optical Theorem and Cutkosky Rules

Whenever the photon propagator is in the $t-$ or $u-$ channel, this will be manifestly real and analytic. However, for an $s-$channel process q^2 will be positive and the logarithm can develop a branch cut. This occurs when

$$m^2 - x(1-x)q^2 < 0 \tag{616}$$

In other words, the branch cut *begins* where $q^2 = 4m^2$, when we are at threshold for production of an electron-positron pair.

It is interesting to calculate the imaginary part of $\hat{\Pi}_2$ for $q^2 > 4m^2$. Recall that, for real X,

$$\mathscr{I}\left(\log(-X + i\epsilon)\right) = +\pi \tag{617}$$

while

$$\mathscr{I}\left(\log(-X - i\epsilon)\right) = -\pi \tag{618}$$

so that going once around the complex X plane results in a change in the imaginary part by 2π. This is one of the signature [mathematical] properties of the logarithm function in terms of complex analysis. In other words, the $i\epsilon$ prescription matters.

Using this property of logs in the complex plane, let us write down the imaginary part of the Feynman graph

$$\mathscr{I}\left(\Pi_2(q^2) - \Pi_2(0)\right)\Big|_{q^2 \pm i\epsilon} = \frac{-2\alpha}{\pi}(\pm\pi)\int_{x_-}^{x_+} dx x(1-x) \tag{619}$$

where $x_\pm$ are the solutions to the zero discriminant equation for $q^2 \geq 4m^2$. Clearly,

$$x_\pm = \frac{1}{2} \pm \frac{1}{2}\beta \quad \text{where} \quad \beta := \sqrt{1 - \frac{4m^2}{q^2}} \tag{620}$$

Then

$$\begin{aligned}\mathscr{I}\left(\Pi_2(q^2 \pm i\epsilon) - \Pi_2(0)\right) &= \mp 2\alpha \int_{-\beta/2}^{+\beta/2} dy\left(\frac{1}{4} - y^2\right) \\ &= \mp\frac{\alpha}{3}\sqrt{1 - \frac{4m^2}{q^2}}\left(1 + \frac{2m^2}{q^2}\right)\end{aligned} \tag{621}$$

where $y = x - 1/2$ and $q^2 > 4m^2$. The amazing thing, as Peskin and Schroeder point out in Ch7.5, is that this is precisely the cross section for production of an electron-positron pair! More on this shortly, but first let us pull out how this physics of $\Pi(q^2) - \Pi(0)$ might change what we mean by the EM interaction strength.

Notice that in a nonrelativistic limit $|q^2| \ll m^2$ so that

$$\begin{aligned}\Pi_2(q^2) - \Pi_2(0) &\to \frac{2\alpha}{\pi}\int_0^1 dx x(1-x)\frac{q^2}{m^2} + \dots \\ &= \frac{2\alpha}{\pi}\left(\frac{q^2}{30m^2} + \mathscr{O}(q^4)\right)\end{aligned} \tag{622}$$

so that in the low-q^2 limit

$$\hat{\Pi}_2(q^2) \to \frac{\alpha}{15\pi}\frac{q^2}{m^2} + \mathscr{O}(q^4) \tag{623}$$

We can extract the position-space potential from this via Fourier Transform:

$$\begin{aligned} V(\vec{x} &= \int \frac{d^3q}{(2\pi)^3} e^{i\vec{q}\cdot\vec{x}} \frac{-e^2}{|\vec{q}|^2[1-\hat{\Pi}_2(-|\vec{q}|^2]} \\ &\to \frac{\alpha}{r} - \frac{4\alpha^2}{15m^2}\delta^3(\vec{x}) \end{aligned} \tag{624}$$

Clearly, this correction to the Coulomb potential is important only at short distances. For the H atom, the only wavefunction which has support this far in is the s-wave: $l = 0$. Accordingly, its energy shift coming from 1-loop QED physics is

$$\Delta E = \int d^3x |\psi(\vec{x})|^2 \left(-\frac{4\alpha}{15m^2}\delta^3(\vec{x})\right) = -\frac{4\alpha}{15m^2}|\psi(0)|^2 \tag{625}$$

This is one piece of the famous *Lamb Shift.*

We can actually do better on approximating the effect of the 1-loop QED correction. Consider the integral

$$V(\vec{x}) = \frac{ie^2}{(2\pi)^2 r} 2 \int_{-\infty}^{+\infty} dQ \frac{Qe^{iQr}}{Q^2+\mu^2}\left[1+\hat{\Pi}_2(q^2)\right] \tag{626}$$

where we defined $|\vec{q}| := Q$ and inserted a fictitious mass μ for the photon to regulate the IR divergence lurking within. What features does our integrand have in the complex plane? Well, it has a branch cut up the imaginary axis starting at $2im$, and it has a pole at $i\mu$. The physical interpretation of these is that the Coulomb potential arises from the pole while QED 1-loop corrections to the Coulomb potential arise from the branch cut.

To evaluate the integral, we push the contour upward. Since the real part of the integrand is the same on both sides of the cut, the only contribution to the integral must come from the imaginary part of $\hat{\Pi}_2$. Defining $q = -iQ$, we have

$$\begin{aligned} \delta V(r) &= \frac{-e^2}{(2\pi)^2 r} 2 \int_{-\infty}^{+\infty} \frac{e^{-qr}}{r} \mathscr{I}(\hat{\Pi}_2(q^2 - i\epsilon)) \\ &= -\frac{\alpha}{r}\frac{2}{\pi}\int_{2m}^{+\infty} dq \frac{e^{-qr}}{q}\frac{\alpha}{3}\sqrt{1-\frac{4m^2}{q^2}}\left(1+\frac{2m^2}{q^2}\right) \end{aligned} \tag{627}$$

At long distances, i.e. when $r \gg 1/m$, this integral is dominated by the region $q \simeq 2m$. Approximating the integrand in this region and substituting $t = q - 2m$ we find

$$\delta V(r) = -\frac{\alpha}{r}\frac{2}{\pi}\int_0^{\infty} dt \frac{e^{-(t+2m)r}}{2m}\frac{\alpha}{3}\sqrt{\frac{t}{m}\frac{3}{2}} + \mathscr{O}(t) \tag{628}$$

and so

$$\delta V(r) \simeq -\frac{\alpha}{r}\frac{\alpha}{4\sqrt{\pi}}\frac{e^{-2mr}}{(mr)^{3/2}} \tag{629}$$

so that the 1-loop corrected potential is

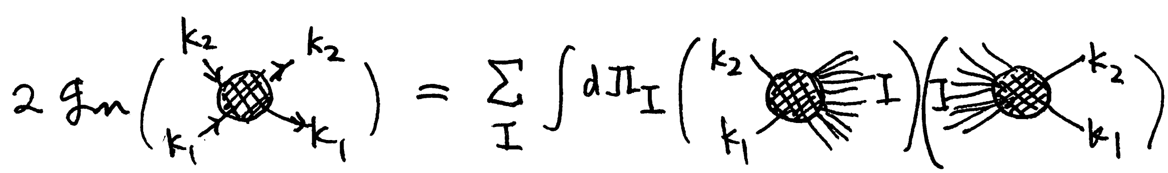
$$V(r) = -\frac{\alpha}{r}\left(1 + \frac{\alpha}{4\sqrt{\pi}}\frac{e^{-2mr}}{(mr)^{3/2}} + \dots\right) \tag{630}$$

Note that the range of the correction term is limited: it leaks out to the electron Compton wavelength $1/m$. On this scale (about a picometre), H atom wavefunctions are approximately constant, amd so the delta function approximation was actually a pretty decent approximation. The radiative correction piece above is known as the *Uehling potential.*

As suggested earlier in the classroom discussion of the running of α_{QED}, the radiative correction can be interpreted as being due to screening. At distances $r \geq 1/m$ virtual e- e+ pairs make the vacuum a dielectric in which the apparent charge (what we measure) is less than the true charge. At smaller distances we begin to penetrate the cloud of virtual dipoles and see more of the bare charge. Unsurprisingly, this phemenon is known as Vacuum Polarization.

Now let us return to thinking about real and imaginary parts. We found that for $q^2 > 4m^2$ the vacuum polarization picked up an imaginary part, which was related physically to the cross section for production of an e- e+ pair.

More generally, *the imaginary part of a forward scattering amplitude arises from a sum of contributions from all possible intermediate states.* This is known as the *Optical Theorem.*



Where does the Optical Theorem come from? Unitarity of the S-matrix. Let us now see how this arises. Consider our friend the S-matrix S and its friend the transfer matrix T:

$$S^\dagger S = 1 \quad \text{where} \quad S = 1 + iT \tag{631}$$

Then we have

$$-i(T - T^\dagger) = T^\dagger T \tag{632}$$

Next, we imagine sandwiching this equation between physical two-particle states $|p_1, p_2\rangle$ and $|k_1, k_2\rangle$. On the RHS we also imagine inserting a complete set of intermediate states, as well. Then the unitarity equation for the transfer matrix becomes

$$\begin{aligned}
&-i\left[\mathscr{M}(k_1k_2 \to p_1p_2) - \mathscr{M}^*(p_1p_2 \to k_1k_2)\right] \\
&= \sum_n \left(\prod_{i=1}^n \frac{d^3q_i}{(2\pi)^3}\frac{1}{2E_i}\right)\mathscr{M}^*(p_1p_2 \to \{q_i\})\mathscr{M}(k_1k_2 \to \{q_i\})\times \\
&\times(2\pi)^d\delta^d(k_1 + k_2 - \sum_i q_i)
\end{aligned} \tag{633}$$

where we have factored off an overall momentum conserving delta function.

Our assumption here was that initial and final states are two-particle asymptotic states. They could just as well have been one- or multi-particle states. Schematically, then, unitarity of the S-matrix implies that

$$-i\left[\mathscr{M}(a\to b)-\mathscr{M}^*(b\to a)\right]=\sum_I\int d\Pi_I\mathscr{M}^*(b\to I)\mathscr{M}(a\to I) \tag{634}$$

For the case of two-particle states, and for the case of forward scattering where $p_i=k_i$ gives

$$\mathscr{I}(\mathscr{M}(k_1k_2\to k_1k_2)=2E_{\rm CM}p_{\rm CM}\sigma_{\rm tot}(k_1k_2\to \text{anything}) \tag{635}$$

where we made use of PS(4.79) for Lorentz-invariant phase space measure.

For Feynman Diagrams, the apperance of an imaginary part always requires a branch cut singularity in the complex (momentum) plane. We saw this quite explicitly at one loop by recognizing the signature property of logarithms in the complex plane. The rules beyond one loop are more complicated than our one-loop example. The awesome thing is that J. Cutkosky proved it generally, and his relations go by the name of Cutkosky Rules (1960).

- Cut through the Feynman diagram in all possible ways such that the cut propagators can be simultaneously put on-shell.
- For each cut, replace each propagator like $1/(p^2-m^2+i\epsilon)$ by $2\pi i\delta(p^2-m^2)$, then do loop integrals.
- Sum the contributions of all possible cuts.

Using these cutting rules, the Optical Theorem can be proven to all orders in perturbation theory.

10.8 Counterterms and the QED beta function

So... how do we actually *renormalize* QED? So far we have seen how to regularize it, using the techniques of dimensional regularization. But what are the systematics of counterterms for QED? Let us start our exposition here with a reminder of the three primitively divergent Feynman graphs of QED:

$$\begin{aligned}\Sigma(p) &= \frac{e^2}{8\pi\varepsilon}(-\not{p}+4m)+\text{finite}\\ \Lambda_\mu(p,p',q) &= \frac{e^2}{8\pi^2\varepsilon}\gamma_\mu+\text{finite}\\ \Pi_{\mu\nu}(k) &= \frac{e^2}{6\pi^2\varepsilon}\left(k_\mu k_\nu-\eta_{\mu\nu}k^2\right)+\text{finite}\end{aligned} \tag{636}$$

Notice that these divergent parts satisfy the Ward Identity - as they must, by gauge invariance.

Consider first the one-loop corrected photon propagator. This involved

$$\Pi_{\mu\nu}(k) = \frac{e^2}{6\pi^2}(k_\mu k_\nu-\eta_{\mu\nu}k^2)\left\{\frac{1}{\varepsilon}-\frac{\gamma}{2}\right.$$

$$
\begin{aligned}
&+3\int_0^1 dx x(1-x)\log\left(\frac{4\pi\mu^2}{-m^2+k^2x(1-x)}\right)+\mathscr{O}(\varepsilon)\Big\} \\
\to\;& \frac{e^2}{6\pi^2}(k_\mu k_\nu - \eta_{\mu\nu}k^2)\left\{\frac{1}{\varepsilon}+\frac{k^2}{10m^2}+\ldots\right\} \qquad (637)
\end{aligned}
$$

Therefore, combining the 1-loop and tree pieces gives for the photon propagator

$$
\begin{aligned}
iD'_{\mu\nu} &= \frac{-i\eta_{\mu\nu}}{k^2} - \frac{e^2}{6\pi^2}\frac{i\eta_{\mu\nu}}{k^2}\left[(k^\alpha k^\beta - \eta^{\alpha,\beta}k^2)\left(\frac{1}{\varepsilon}+\frac{k^2}{10m^2}\right)\right]\frac{\eta_{\beta\nu}}{k^2}+\ldots \\
&= -i\frac{\eta_{\mu\nu}}{k^2}\left(1+\frac{e^2}{6\pi^2\varepsilon}+\frac{e^2}{60\pi^2}\frac{k^2}{m^2}\right) - \frac{e^2}{6\pi^2\varepsilon}\frac{k_\mu k_\nu}{k^4}+\ldots \qquad (638)
\end{aligned}
$$

Notice that this corrected propagator is *not* in Feynman gauge. The reason is the term proportional to $k_\mu k_\nu$. However, as we saw earlier, this does not affect physical quantities which are gauge-invariant, because of the Ward Identity. The details of renormalization do depend sensitively on both the gauge and the renormalization scheme.

The infinite $1/\varepsilon$ pieces must be removed by adding counterterms to the original Lagrangian. For QED, in the gauge sector we had in Feynman gauge

$$
\mathscr{L}_2 = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2}(\partial\cdot A)^2 = \frac{1}{2}A^\mu(\eta_{\mu\nu}\partial^2)A^\nu \qquad (639)
$$

So the counterterm we need at one loop is

$$
\Delta\mathscr{L}_2^{\mathrm{CT}} = -\frac{C}{4}F^{\mu\nu}F_{\mu\nu} - \frac{E}{2}(\partial\cdot A)^2 \qquad (640)
$$

Note that C and E will be different, because of our observation above that the one-loop corrected propagator is not in Feynman gauge even though the tree level propagator was. So our bare Lagrangian becomes

$$
\mathscr{L}_{2\,B} = -\frac{(1+C)}{4}F^{\mu\nu}F_{\mu\nu} - \frac{(1+E)}{2}(\partial_\lambda A^\lambda)^2 \qquad (641)
$$

We can rewrite this as

$$
\begin{aligned}
\mathscr{L}_{2\,B} &= -\frac{(1+C)}{4}F^{\mu\nu}F_{\mu\nu} + \text{gauge terms} \\
&= -\frac{Z_3}{4}F^{\mu\nu}F_{\mu\nu} + \text{gauge terms} \qquad (642)
\end{aligned}
$$

where

$$
Z_3 = 1 - \frac{e^2}{6\pi^2\varepsilon} \qquad (643)
$$

In this way, we obtain a finite 1-loop propagator for our friendly photon. Most importantly, its mass is still exactly zero. Yay!

From earlier in this section on 1-loop QED, we had for the propagator

$$
iD'_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{k^2[1-\Pi(k^2)]} + \text{gauge terms}
$$

$$
\begin{aligned}
&= \frac{-i\eta_{\mu\nu}}{k^2[1-\Pi_{\text{finite}}(k^2)+e^2/(6\pi^2\varepsilon)]} + \text{gauge terms} \\
&= -\frac{iZ_3\eta_{\mu\nu}}{k^1[1-\Pi_{\text{finite}}(k^2)]} + \text{gauge terms}
\end{aligned}
\tag{644}
$$

where we used our earlier definition $Z_3 = 1/[1-\Pi(0)]$. From the definition of a propagator, we know it is the vacuum-to-vacuum transition amplitude involving a time-ordered product of two field operators. Using this intuition, we see that the counterterm Lagrangian therefore motivates the definition of the *bare* gauge field as

$$
A_B^\mu = \sqrt{Z_3}A_\mu \tag{645}
$$

It follows immediately that the renormalized propagator is

$$
i\tilde{D}'_{\mu\nu} = \frac{-i\eta_{\mu\nu}}{k^2[1-\Pi_{\text{finite}}(k^2)]} + \text{gauge terms} \tag{646}
$$

Now let us turn to the electron self-energy.

We found earlier in this section that the electron self-energy at 1-loop modifies the electron's inverse propagator. Our tree+1-loop expression is

$$
\begin{aligned}
\Gamma^{(2)}(p) &= S'_F(p)^{-1} = S_F(p)^{-1} - \Sigma(p) \\
&= (\not{p}-m) - \frac{e^2}{8\pi^2\varepsilon}(-\not{p}+4m) + \text{finite} \\
&= \not{p}\left(1+\frac{e^2}{8\pi^2\varepsilon}\right) - m\left(1+\frac{e^2}{2\pi^2\varepsilon}\right) + \text{finite}
\end{aligned}
\tag{647}
$$

Because the coefficients of $\not{p}$ and m are not equal, a single counterterm will not suffice to kill the UV infinities. We will need one counterterm for the overall magnitude of the propagator, contributing to electron wavefunction renormalization, and a separate one for the electron mass. So we add

$$
\mathscr{L}_1^{(\text{CT})} = iB\bar{\psi}\not{\partial}\psi - A\bar{\psi}\psi \tag{648}
$$

giving the total – bare – Lagrangian for the electron sector alone

$$
\mathscr{L}_{1\,B} = i(1+B)\bar{\psi}\not{\partial}\psi - (m+A)\bar{\psi}\psi \tag{649}
$$

So our counterterms are:

Note that, as you can easily check, adding the interaction $iB\not{p}$ modifies a propagator from $i/\not{p}$ to $i/[(1+B)\not{p}]$. Hence

$$
\frac{e^2}{8\pi^2\varepsilon}(-\not{p}+4m) + A - B\not{p} = \text{finite} \tag{650}
$$

as long as we set

$$
A = -\frac{e^2}{2\pi^2\varepsilon} \quad \text{and} \quad B = -\frac{i^2}{8\pi^2\varepsilon} \tag{651}
$$

The coefficient in front of the canonically normalized electron kinetic energy is

$$Z_2 = 1 + B = 1 - \frac{e^2}{8\pi^2\varepsilon} \tag{652}$$

This enables us to define the 'bare' wave function by

$$\psi_B = \sqrt{Z_2}\psi \tag{653}$$

in terms of which we can write the bare electron Lagrangian:

$$\mathscr{L}_{1\,B} - i\bar{\psi}_B \partial\!\!\!/ \psi_B - m_B \bar{\psi}_B \psi_B \tag{654}$$

where the bare mass is defined by

$$m_B = Z_2^{-1}(m + A) = \left(1 - \frac{e^2}{2\pi^2\varepsilon}\right) m \left(1 + \frac{e^2}{8\pi^2\varepsilon}\right) = m\left(1 - \frac{3e^2}{8\pi^2\varepsilon}\right) = m + \delta m \tag{655}$$

at one loop order. Our philosophy here is that ψ is the physical electron field, while ψ_B is the bare electron field which appears in the Lagrangian from which vertices for calculation are taken.

Accordingly, our renormalized inverse electron propagator at one loop becomes

$$\Gamma^{(2)}(p) = (p\!\!\!/ - m) - [\Sigma(p) + A - Bp\!\!\!/] = p\!\!\!/ - m + \text{finite} \tag{656}$$

We turn, finally, to the vertex function and its divergent part Λ_μ. We had previously that

$$\Lambda_\mu(p, q, p' = p + q) = \frac{e^2}{8\pi^2\varepsilon}\gamma_\mu + \text{finite} \tag{657}$$

The effect of this singular piece can be eliminated by adding a counterterm to the Lagrangian of the form

$$\Delta\mathscr{L}_3^{(\text{CT})} = -De\mu^{2-D/2}\bar{\psi}A\!\!\!/\psi \tag{658}$$

where

$$D = -\frac{e^2}{8\pi^2\varepsilon} \tag{659}$$

The Bare Lagrangian for the QED vertex at tree+1-loop is therefore

$$\mathscr{L}_{3\,B} = -(1 + D)e\mu^{\varepsilon/2}A^\mu\bar{\psi}\gamma_\mu\psi := -Z_1 e\mu^{\varepsilon/2}A^\mu\bar{\psi}\gamma_\mu\psi \tag{660}$$

with

$$Z_1 = 1 - \frac{e^2}{8\pi^2\varepsilon} \tag{661}$$

We have now completed the business of regularizing and renormalizing QED to one-loop. Our total bare Lagrangian at this one-loop order is

$$\mathscr{L}_B = iZ_2\bar{\psi}\gamma^\mu\partial_\mu\psi - (m+A)\bar{\psi}\psi - Z_1 e\mu^{\varepsilon/2}A^\mu\bar{\psi}\gamma_\mu\psi - Z_3\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 + \text{gauge terms} \tag{662}$$

with

$$Z_1 = Z_2 = 1 - \frac{e^2}{8\pi^2\varepsilon} \qquad Z_3 = 1 - \frac{e^2}{6\pi^2\varepsilon} \qquad A = -m\frac{e^2}{2\pi^2\varepsilon} \tag{663}$$

This Lagrangian gives – to one loop order in perturbation theory – finite self-energies and vertex, where e and m stand for physical quantities which are experimentally measurable in the lab. The bare charge is

$$e_B = e\mu^{\varepsilon/2}\frac{Z_1}{Z_2\sqrt{Z_3}} = e\mu^{\varepsilon/2}Z_3^{-1/2} \tag{664}$$

The total bare Lagrangian is, counting all sectors,

$$\mathscr{L}_B = i\bar{\psi}_B\partial\!\!\!/\psi_B - m_B\bar{\psi}_B\psi_B - e_BA_B^\mu\bar{\psi}_B\gamma_\mu\psi_B - \frac{1}{4}\left(\partial_\mu A_{B\,\nu} - \partial_\nu A_{B\,\mu}\right)^2 \tag{665}$$

From our relationships between bare and renormalized quantities, we obtained

$$e_B = e\mu^{\varepsilon/2}\left(1 + \frac{e^2}{12\pi^2\varepsilon}\right) \tag{666}$$

In the limit $\varepsilon \to 0$, the bare charge e_B is independent of μ, but the physical charge e does depend on μ as we now show. Differentiating the above equation w.r.t. μ gives, to one loop order in QED perturbation theory,

$$\mu\frac{\partial e}{\partial\mu} = -e\frac{\varepsilon}{2} + \frac{e^3}{12\pi^2} \tag{667}$$

Letting $\varepsilon \to 0$ then gives

$$\beta(\mu) = \mu\frac{\partial e}{\partial\mu} = \frac{e^3}{12\pi^2} \tag{668}$$

So, like $\lambda\phi^4$ scalar field theory, QED has a positive beta function. This differential equation for the QED coupling $e(\mu)$ at one-loop order has the solution

$$e^2(\mu) = \frac{e^2(\mu_0)}{1 - [e^2(\mu_0)/(6\pi^2)]\log(\mu/\mu_0)} \tag{669}$$

From this we can see very explicitly the increase of e with μ. If we stare hard at the formula, we can also see that it possesses a sickness known as the *Landau pole*: the physical coupling blows up at $\mu = \mu_*$ when

$$\mu_* = \mu_0\exp\left(\frac{6\pi^2}{e^2(\mu_0)}\right) \tag{670}$$

This is ameliorated by higher loop corrections, but QED is still a sick theory in the sense that it becomes strongly coupled in the UV. This indicates that it must be embedded in a larger theory. Nonabelian gauge theories, by contrast, have negative beta functions, which is something you will discover for yourselves in your Final Project.

11 An introduction to chiral anomalies

One of the central tenets of functional quantization is keeping symmetries of the theory manifest in the Lagrangian of the theory and all physics calculated therefrom. Sometimes, however, there can be a hiccup that arises on doing loop-level calculations in any given QFT. This hiccup takes the form of quantum breaking of symmetries of the classical action. The resulting pain in the arse is known as an *anomaly*.

There are many ways to see how the Adler-Bell-Jackiw anomaly arises. One derivation focuses on the behaviour of the measure in the fermionic Feynman Path Integral. Since this fits best with our functional quantization exposition this semester, we will reveal the anomaly using the FPI.

11.1 Anomalies in Path Integral Quantization

First we do a 'quick-and-dirty' attempt to get the right answer. It will turn out that this result is incorrect and will need to be refined.

A prototypical example is the axial current. Consider a species of fermion ψ with Feynman Path Integral

$$Z = \int \mathscr{D}\bar{\psi}\mathscr{D}\psi \exp\left[i\int d^4x\left(\bar{\psi}i\not{D}\psi\right)\right] \tag{671}$$

Suppose that we perform a chiral rotation, changing variables to

$$\begin{aligned}\psi(x) &\to \psi'(x) = (1 + i\alpha(x)\gamma_5)\,\psi(x)\\ \bar{\psi}(x) &\to \bar{\psi}'(x) = \bar{\psi}(x)\,(1 + \alpha(x)\gamma_5)\end{aligned} \tag{672}$$

Now, a global chiral rotation – one with *constant* α – is a symmetry of the tree-level Lagrangian. Therefore, the only new terms generated in $\mathscr{L}$ by doing the above chiral transformation are proportional to $\partial\alpha(x)$. Specifically,

$$\begin{aligned}\int d^4x\bar{\psi}'(i\not{D})\psi' &= \int d^4x\left[\bar{\psi}i\not{D}\psi - \partial_\mu\alpha(x)\bar{\psi}\gamma^\mu\gamma_5\psi\right]\\ &= \int d^4x\left[\bar{\psi}i\not{D}\psi + \alpha(x)\partial_\mu\left(\bar{\psi}\gamma^\mu\gamma_5\psi\right)\right]\end{aligned} \tag{673}$$

where the last line above follows using integration by parts. Finally, varying the Lagrangian w.r.t. $\alpha(x)$, we obtain the chiral current conservation equation

$$\partial_\mu j^{\mu 5} = 0 \tag{674}$$

So is this really true? Is chiral symmetry preserved at loop level for an arbitrary QFT?

Oops. We made one huge assumption in our method above:

$$\mathscr{D}\bar{\psi}'\mathscr{D}\psi' = \mathscr{D}\bar{\psi}\mathscr{D}\psi \tag{675}$$

Unfortunately, it turns out that we have been entirely too cavalier here. We just assumed that there was no Jacobian involved. We should now take a couple of steps back and take time to calculate how the fermionic measure actually transforms.

K.Fujikawa in 1979 showed what goes wrong quantum mechanically with assuming chiral current conservation. Following his analysis (and the exposition in Peskin and Schroeder!), we now switch our close focus to the behaviour of the fermion measure in the FPI.

To do a proper job of defining the FPI measure, let us expand our fermion field $\psi(x)$ in a basis of eigenstates ϕ_m of the operator $\not{D}$. In particular, let us define L and R eigenvectors via

$$\begin{aligned}(i\not{D})\phi_m &= \lambda_m\phi_m \\ \hat{\phi}_m(i\not{D}) &= -iD_\mu\hat{\phi}_m\gamma^\mu = \lambda_\mu\hat{\phi}_m\end{aligned} \tag{676}$$

When the background A_μ gauge field is zero, these eigenstates are just the regular Diract wavefunctions of definite momentum – animals with which you are intimately familiar as a result of QFT1. The eigenvalues λ_m obey

$$\lambda_m^2 = k^2 = (k^0)^2 - |\vec{k}|^2 \tag{677}$$

But for a *fixed* background gauge field, this is also (!) the asymptotic form of the eigenvalues for large k. For anyone interested in deeper level of detail, Tom Banks' textbook "Modern Quantum Field Theory" section 8.8 contains a more detailed analysis than we attempt here. From now on this asymptotic form of the eigenvalues will be the driver.

We write in this basis

$$\begin{aligned}\psi(x) &= \sum_m a_m\phi_m(x) \\ \bar{\psi}(x) &= \sum_m \hat{a}_m\hat{\phi}_m(x)\end{aligned} \tag{678}$$

where $a_m, \hat{a}_m$ are *anticommuting* coeffficients and $\phi_m, \hat{\phi}_m$ are c-number (commuting) wavefunctions.

Then our functional measure over $\psi, \bar{\psi}$ is

$$\mathscr{D}\psi\mathscr{D}\bar{\psi} = \prod_m da_m d\hat{a}_m \tag{679}$$

Suppose that we perform a chiral rotation on our fermion field. How does the measure of the transformed fermion field compare to the original? We have

$$\psi'(x) = \{1 + i\alpha(x)\gamma_5\}\,\psi(x) \tag{680}$$

The expansion coefficients of ψ and ψ' are related by an infinitesimal transformation $(1+C)$, as follows:

$$\begin{aligned}a'_m &= \sum_n \int d^4x\phi_m^\dagger\left(1 + i\alpha(x)\gamma_5\right)\phi_m(x)a_n \\ &= \sum_n \left(\delta_{mn} + C_{mn}\right)a_n\end{aligned} \tag{681}$$

The above tells us that our Jacobian of the chiral transformation is $J = (1 + C)$.

Now comes the really crucial bit. Recall that, by properties of Grassmann integration for anticommuting fields,

$$\mathscr{D}\psi'\mathscr{D}\bar{\psi}' = \frac{1}{J^2}\mathscr{D}\psi\mathscr{D}\bar{\psi} \tag{682}$$

Note that J^2 being in the *denominator* is not a typo: it really is there, by properties of Grassmann integration.

The Jacobian is

$$\begin{aligned} J &= \det(1+C) = \exp\left(\mathrm{Tr}\log(1+C)\right) \\ &= \exp\left(\sum_n C_{nn} + \mathscr{O}(C^2)\right) \end{aligned} \tag{683}$$

and we can ignore higher order terms as C is infinitesimal. So

$$\log J = i\int d^4x\alpha(x)\sum_n \hat{\phi}_n(x)\gamma_5\phi_n(x) \tag{684}$$

It might seem at first glance that this is $tr(\gamma_5) = 0$. But this turns out be an illusion: we have to be careful to regularize the sum over eigenstates in a gauge-invariant way rather than just operating by the seat of the pants.

As with any loop-level Feynman diagram, we can perform any one of a number of regularization methods. Here we make the 'natural choice' (PS) and put in a Gaussian style regulator:

$$\sum_n \phi_n^\dagger(x)\gamma_5\phi_n(x) = \lim_{M\to\infty}\sum_n \phi_n^\dagger(x)\gamma_5\phi_n(x)\exp\left(\frac{\lambda_n^2}{M^2}\right) \tag{685}$$

Note: because we work in a mostly-minus signature, the above choice of sign in the Gaussian factor is indeed correct:

$$k^2 = (k^0)^2 - |\vec{k}|^2 \tag{686}$$

and upon Wick rotation the above Gaussian damping factor becomes

$$\exp\left(-(\lambda_E)_n^2/M^2\right) \tag{687}$$

as expected.

In operator form, our expression becomes

$$\begin{aligned} \lim_{M\to\infty} \quad \textstyle\sum_n \quad & \phi_n^\dagger(x)\gamma_5\exp\left(\frac{(i\not{D})^2}{M^2}\right)\phi_n(x) \\ = \quad & \lim_{M\to\infty}\langle x|\mathrm{Tr}\left[\gamma_5\exp\left(\frac{(i\not{D})^2}{M^2}\right)\right]|x\rangle \end{aligned} \tag{688}$$

where the trace indicated is over Dirac fermion indices.

Our next step is to evaluate $(i\not{D})^2$. We know that

$$\begin{aligned} (i\not{D})^2 &= -(\not{D})^2 = -\gamma^\mu D_\mu\gamma^\nu D_\nu \\ &= -\gamma^\mu\gamma^\nu D_\mu D_\nu \end{aligned} \tag{689}$$

Now, we can always expand a two-index tensor in terms of its symmetric and antisymmetric parts:

$$\gamma^\mu \gamma^\nu = \gamma^{(\mu}\gamma^{\nu)} + \gamma^{[\mu}\gamma^{\nu]} \qquad (690)$$

The first term on the RHS is just the Minkowski metric, by the fundamental anticommutation relations for gamma-matrices. Defining

$$\gamma^{\mu\nu} \equiv \frac{1}{2}\left[\gamma^\mu, \gamma^\nu\right] \qquad (691)$$

we have

$$\gamma^\mu \gamma^\nu = \eta^{\mu\nu} + \gamma^{\mu\nu} \qquad (692)$$

Therefore, we have for the square of the Dirac operator

$$\begin{aligned}(i\not{D})^2 &= -\left(\eta^{\mu\nu} + \gamma^{\mu\nu}\right) D_\mu D_\nu \\ &= -D^2 + -\gamma^{\mu\nu} D_\mu D_\nu \qquad (693)\end{aligned}$$

Clearly, the term $\gamma^{\mu\nu} D_\mu D_\nu$ can be nonzero only if $[D_\mu, D_\nu] \neq 0$. We know from our earlier discussion of non-Abelian gauge field theories that, acting on fermions,

$$[D_\mu, D_\nu] = -igF_{\mu\nu} \qquad (694)$$

(You can check this explicitly by using $D_\mu = \partial_\mu - igA_\mu$.) Therefore

$$\begin{aligned}(i\not{D})^2 &= -D^2 - \gamma^{\mu\nu}\left(\frac{-ig}{2}F_{\mu\nu}\right) \\ &= -D^2 + \frac{ig}{2}\gamma^{\mu\nu}F^B_{\mu\nu}t^B \\ &= -D^2 - \frac{g}{2}\left(i\gamma^{\mu\nu}\right)F_{\mu\nu} \qquad (695)\end{aligned}$$

As indicated previously, as we take our regulator $M \to \infty$, we can focus our attention on the asymptotic part of the spectrum. We approximate k as large, and expand in powers of the gauge field A_μ. But what kind of term would arise? Properties of γ-matrices in $D = 4$ tell us that we will get a result tracing γ_5 against $e^{(i\not{D})^2/M^2}$ if we have *four* γ-matrices beside the γ_5. This intuition allows us to anticipate the result to be a term of order $(i\gamma : F)^2$, and other terms will be subdominant to it. We find

$$\begin{aligned}&\lim_{M\to\infty} \langle x| \ \mathrm{Tr}\left\{\exp\left[-D^2 - \frac{ig}{2}(i\gamma^{\mu\nu})F_{\mu\nu}\right]\right\}|x\rangle \\ &= \lim_{M\to\infty} \mathrm{Tr}\left\{\gamma_5 \frac{1}{2!}\left[\frac{g}{2M^2}\left(i\gamma^{\mu\nu}F_{\mu\nu}\right)\right]^2\right\}\langle x|e^{-\partial^2/M^2}|x\rangle \qquad (696)\end{aligned}$$

Now,

$$\begin{aligned}\langle x|e^{-\partial^2/M^2}|x\rangle &= \lim_{x\to y}\int \frac{d^4k}{(2\pi)^4}\exp\left(-ik\cdot(x-y)\right)\exp\left(\frac{k^2}{M^2}\right) \\ &= i\int \frac{d^4k_E}{(2\pi)^4}\exp\left(-\frac{k_E^2}{M^2}\right)\end{aligned}$$

$$= i\frac{M^4}{16\pi^2} \tag{697}$$

Therefore, we have

$$\lim_{M\to\infty}\left\{\frac{1}{8}\frac{-ig^2}{16\pi^2}M^4\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\gamma^\lambda\gamma^\sigma\left(\frac{1}{M^2}\right)^2F_{\mu\nu}F_{\lambda\sigma}\right]\right\}$$
$$= \frac{-g^2}{32\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma} \tag{698}$$

In other words,

$$J = \exp\left\{-i\int d^4x\alpha(x)\frac{g^2}{32\pi^2}\varepsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}\right\} \tag{699}$$

Look carefully at what just happened! After our change of variables according to a chiral rotation

$$(\psi,\bar{\psi}) \to (\psi',\bar{\psi}') \tag{700}$$

our Feynman Path Integral picked up a factor

$$Z = \int \mathscr{D}\psi\mathscr{D}\bar{\psi}\exp\left(i\int d^4x\left\{\bar{\psi}(i\not{D})\psi + \alpha(x)\left[\partial_\mu j^{\mu 5} + \frac{g^2}{16\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma}\right]\right\}\right) \tag{701}$$

Since this has to be true for arbitrary $\alpha(x)$, the loop-modified equation for nonconservation of $j^{\mu 5}$ is

$$\partial_\mu j^{\mu 5} = -\frac{g^2}{16\pi^2}\epsilon^{\mu\nu\lambda\sigma}F_{\mu\nu}F_{\lambda\sigma} \tag{702}$$

For Yang-Mills theories rather than QED, we would add a trace in front of the epsilon pseudotensor.

What physics message should we draw from this? Well, since the chiral current is not conserved in a background gauge field, the number of fermions minus the number of antifermions will not be conserved.

Note: this derivation we just completed generalizes readily to any *even* dimension of spacetime. The functional derivation always picks out the term in the expansion of $\exp\left(i\gamma^{\mu\nu}F_{\mu\nu}\right)$ that has the same dimension, D, as the divergence of the anomalous current. The general expression is

$$\partial_\mu j^{\mu 5} = (-1)^{D/2+1}\frac{2e^{D/2}}{(D/2)!(4\pi)^{D/2}}\epsilon^{\mu_1\mu_2\dots\mu_{D/2}}F_{\mu_1\mu_2}\dots F_{\mu_{D-1}\mu_D} \tag{703}$$

11.2 Triangle anomaly: the Feynman Diagram approach

Consider our friend the Standard Model. Recall our Pauli sigma matrix conventions

$$\begin{aligned}\{\sigma^\mu\} &= \{\mathbb{1},+\vec{\sigma}\}\\ \{\bar{\sigma}^\mu\} &= \{\mathbb{1},-\vec{\sigma}\}\end{aligned} \tag{704}$$

Also, our covariant derivative acting on any given fermion representation r is

$$D_\mu = \mathbb{1}\partial_\mu - igA^A_\mu t^A_r \tag{705}$$

where the generators t_r depend on the fermion representation.

Suppose that we keep right-handed chiral fermions free and minimally couple the left-handed chiral fermions to the gauge field A_μ. Then the Lagrangian becomes

$$\mathscr{L} = \bar{\psi} i \gamma^\mu \left[\mathbb{1} \partial_\mu - i g A_\mu^A t_r^A \left(\frac{\mathbb{1} - \gamma_5}{2} \right) \right] \psi \tag{706}$$

This Lagrangian is invariant under gauge transformations on left-handed fields only:

$$\begin{aligned} \psi &\to \left[\mathbb{1} + i \alpha^A t_r^A \left(\frac{\mathbb{1} - \gamma_5}{2} \right) \right] \psi \\ A_\mu^A &\to A_\mu^A + \frac{1}{g} \partial_\mu \alpha + f^{ABC} A_\mu^B \alpha^C \end{aligned} \tag{707}$$

which are just the usual Yang-Mills transformation laws. The corresponding Noether current is

$$j^{\mu A} = \bar{\psi} \gamma^\mu \left(\frac{\mathbb{1} - \gamma_5}{2} \right) t_r^A \psi \tag{708}$$

The projector

$$\mathscr{P}_- = \frac{\mathbb{1} - \gamma_5}{2} \tag{709}$$

is involved in the Noether current because only left-handed fields contribute to it: they are the only fields that "feel" the gauge symmetry. The right-handers are gauge singlets in this model.

We are now in a position to ask the question: how can we see non-conservation of the chiral Noether current in the context of Feynman diagrams at loop level in perturbation theory? Consider $\partial_\mu j^{\mu A}$; in momentum space this will be proportional to $i q_\mu j^{\mu A}$. Therefore, the Feynman diagrams that are relevant will involve an insertion of $j^{\mu A}$ into a one-loop diagram. In turn, these Feynman diagrams must involve fermions running around the loop, and by the structure of $\psi - \bar{\psi} - A$ vertices, they involve two external gauge boson legs. Diagrammatically:

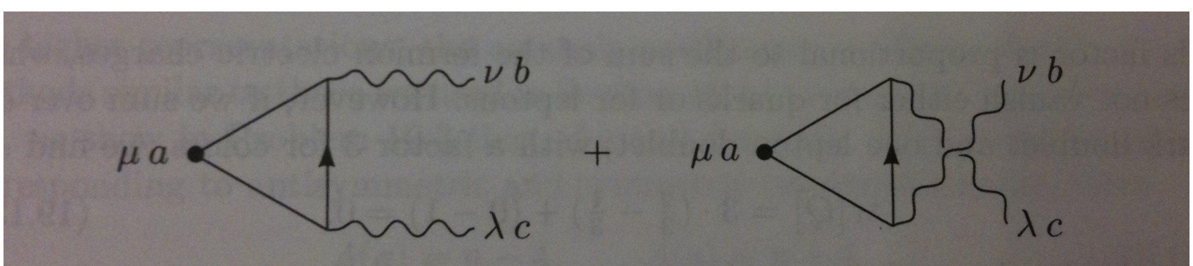


Figure 19.9. Diagrams contributing to the anomaly of a gauge symmetry current in a chiral gauge theory.

From our Path Integral discussion of anomalies, we can anticipate that only the γ_5 part of the chiral Noether current will contribute.

In terms of matrix elements,

$$\int d^4 x e^{-iq\cdot x} \langle p, \nu; k, \lambda | j^{\mu 5}(x) | 0 \rangle = (2\pi)^4 \delta^{(4)}(p + k - q) \epsilon_\nu^*(p) \epsilon_\lambda^*(k) \mathscr{M}^{\mu\nu\lambda}(p, k) \tag{710}$$

where $\mathscr{M}^{\mu\nu\lambda}$ gets contributions from

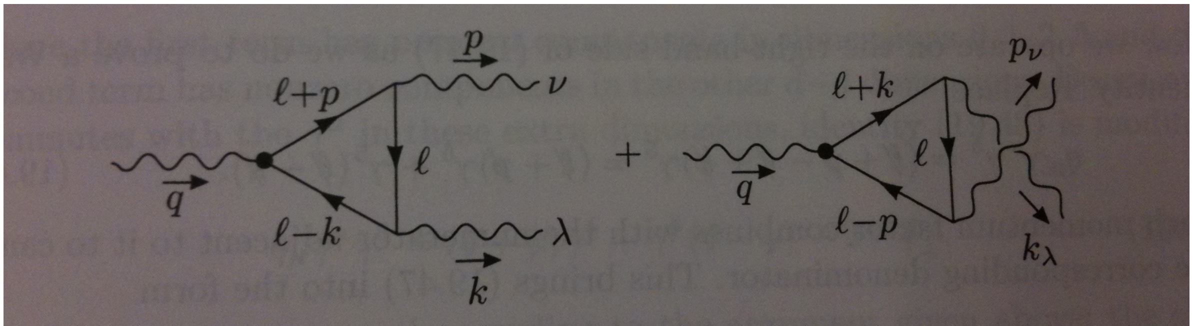


Figure 19.4. Diagrams contributing to the two-photon matrix element of the divergence of the axial vector current.

Let us evaluate the first diagram. Since the anomaly equation involves the divergence of the chiral Noether current, in momentum space this means contracting up the above with iq_μ:

$$Q_1^\mu \equiv (-1)(-ie)^2 \int \frac{d^D l}{(2\pi)^D} \mathrm{Tr}\left[(iq_\mu)\gamma^\mu\gamma_5 \frac{i(\not{l}+\not{k})}{(l-k)^2}\gamma^\lambda \frac{i\not{l}}{l^2}\gamma^\nu \frac{i(\not{l}-\not{p})}{(l+p)^2}\right] \tag{711}$$

The second diagram just has $(p,\nu) \leftrightarrow (k,\lambda)$.

Now

$$\begin{aligned} q_\mu\gamma^\mu\gamma_5 &= \not{q}\gamma_5 \\ &= [\not{k}+\not{p}+(\not{l}-\not{l})]\gamma_5 \\ &= (\not{l}+\not{p})\gamma_5 + \gamma_5(\not{l}-\not{k}) \end{aligned} \tag{712}$$

So iq_μ dotted into the diagram with the insertion of $j^{\mu 5}$ is

$$Q_1 = iq_\mu Q_1^\mu = \int \frac{d^D l}{(2\pi)^D} \mathrm{Tr}\left[\{\not{q}\gamma_5 = (\not{l}+\not{p})\gamma_5 + \gamma_5(\not{l}-\not{k})\}\frac{1}{(\not{l}-\not{k})}\gamma^\lambda \frac{1}{(\not{l})}\gamma^\nu \frac{1}{(\not{l}+\not{p})}\right] \tag{713}$$

Notice that one propagator factor cancels part of the numerator in each of the two terms.

Next, we assume that in dimensional regularization it is fine to gaily anticommute γ_5 through other γ^ν, which yields

$$Q_1 = e^2 \int \frac{d^D l}{(2\pi)D} \mathrm{Tr}\left[+\gamma_5 \frac{1}{(\not{l}-\not{k})}\gamma^\lambda \frac{1}{\not{l}}\gamma^\nu - \gamma_5 \frac{1}{\not{l}}\gamma^\nu \frac{1}{(\not{l}+\not{p})}\gamma^\lambda\right] \tag{714}$$

where we have also made use of cyclicity of the trace in both terms. Notice that we can relabel $l \to l+k$ in the first loop integral. It follows immediately that our result for the first diagram contributing to Q_1 is manifestly antisymmetric under interchange of the labels on the two external gauge boson legs. But the sum of the Feynman graphs must be symmetric under the very same interchange. Therefore, the sum of the two Feynman graphs is zero and the chiral anomaly is apparently *nowhere to be seen!*

So what did we do wrong? Our fatal assumption was regarding the anticommutation behaviour of γ_5. In truth, when we continue away from four dimensions of spacetime, there are surprises in store, arising from the fact that γ_5 is an intrinsically four-dimensional animal. To get at the correct answer, we need to recall that the loop momentum over which we integrate lives in D dimensions of spacetime, not four, so that we can write

$$l = l_{\|} + l_{\perp} \tag{715}$$

where $l_{\|}$ lives in four dimensions and $l_{\perp}$ lives in $D-4$ dimensions. As a consequence, we have the extremely important revision of the familiar behaviour of γ_5:

$$\begin{aligned} \{\gamma^{\|}, \gamma_5\} &= 0 \\ [\gamma^{\perp}, \gamma_5] &= 0 \end{aligned} \tag{716}$$

In other words, γ_5 only anticommutes with the four-dimensional set of γ^μ, and commutes with the rest (because it has no reason to anticommute!).

Therefore, we have

$$\begin{aligned} q_\mu \gamma^\mu \gamma_5 &= \not{q}\gamma_5 \\ &= (\not{k} + \not{p})\gamma_5 \\ &= \not{k}\gamma_5 - \gamma_5 \not{p} \\ &= (\not{k} + \not{l}_{\|})\gamma_5 + \gamma_5(\not{l}_{\|} - \not{p}) \\ &= (\not{k} + \not{l})\gamma_5 - \not{l}_{\perp}\gamma_5 + \gamma_5(\not{l} - \not{p}) - \gamma_5 \not{l}_{\perp} \\ &= (\not{k} + \not{l})\gamma_5 + \gamma_5(\not{l} - \not{p}) - 2\gamma_5 \not{l}_{\perp} \end{aligned} \tag{717}$$

Therefore, the only non-cancelling part of the sum of the two Feynman graphs is

$$iq_\mu (Q_1^\mu + Q_2^\mu) = e^2 \int \frac{d^D l}{(2\pi)^D} \mathrm{Tr}\left[\{-2\gamma_5 \not{l}_{\perp}\} \frac{(\not{l} - \not{k})}{(l-k)^2} \gamma^\lambda \frac{\not{l}}{l^2} \gamma^\nu \frac{(\not{l} + \not{p})}{(l+p)^2}\right] \tag{718}$$

Our next few steps are clear, given our past experience with one-loop diagrams.

- Use Feynman parameters (three of them) to combine the three denominators.
- Shift $l \to l + P$ where $P = xk - yp$.
- In expanding the numerator, retain one factor each of $\gamma^\nu, \gamma^\lambda, \not{p}, \not{k}$ to get a nonzero trace against γ_5.
- Focus on what is left over: one factor of $\not{l}_{\perp}$ and one of $\not{l} = (\not{l}_{\|} + \not{l}_{\perp})$. Note that $\not{l}_{\perp}$ commutes with other gamma matrices in the four-dimensional part. Therefore, our final result is of the form

$$\int \frac{d^D l}{(2\pi)^D} \frac{\not{l}_{\perp} \not{l}_{\perp}}{(l^2 - \Delta)^3} \tag{719}$$

  where

$$\frac{i}{(4\pi)^{D/2}} \frac{(D-4)}{2} \frac{\Gamma(2-D/2)}{\Gamma(3)\Delta^{2-D/2}} \longrightarrow \frac{-i}{32\pi^2} \tag{720}$$

  as $D \to 4$. Notice that we did not need to know details about Δ because the $(D-4)$ factor in the numerator cancelled the logarithmic divergence from the Δ-dependent part.

- Use the symmetry relationship

$$(l\!\!\!/_\perp)^2 = l_\perp^2 = \frac{(D-4)}{D} l^2 \tag{721}$$

which is valid under the shifted-l loop integral.

So, after the dust settles, we are left with

$$\begin{aligned} iq_\mu(Q_1^\mu) &= e^2 \frac{-i}{32\pi^2} \mathrm{Tr}\left[2\gamma_5(-k\!\!\!/)\gamma^\lambda(p\!\!\!/)\gamma^\nu\right] \\ &= \frac{e^2}{4\pi^2} \epsilon^{\alpha\lambda\beta\nu} k_\alpha p_\beta \end{aligned} \tag{722}$$

This term is already symmetric in $(p,\nu) \leftrightarrow (k,\lambda)$, and so addition of $iq_\mu Q_2^\mu$ just doubles the final result. Finally, we have

$$\begin{aligned} \langle p,\nu;k,\lambda|\partial_\mu j^{\mu 5}(0)|0\rangle &= \frac{e^2}{2\pi^2} \epsilon^{\alpha\lambda\beta\nu}(-ip_\alpha)\epsilon_\nu^*(p)(-ik_\beta)\epsilon_\lambda^*(k) \\ &= \frac{e^2}{16\pi^2} \langle p,\nu;k,\lambda|\epsilon^{\alpha\nu\beta\lambda} F_{\alpha\nu} F_{\beta\lambda}(0)|0\rangle \end{aligned} \tag{723}$$

This is nothing other than the ABJ anomaly equation in disguise:

$$\partial_\mu j^{\mu 5}(x) = -\frac{e^2}{16\pi^2} \epsilon^{\alpha\beta\gamma\delta} F_{\alpha\beta} F_{\gamma\delta}(x) \tag{724}$$

If we were doing Yang-Mills gauge theory we would simply put a trace before the field strength tensors in the above expression.

11.3 Anomaly cancellation for chiral gauge field theories

Start with a theory containing no mass terms for fermions - like our friend the Standard Model. In a helicity basis, we can couple (at least classically!) left- and right-handed fermions to gauge fields in different ways.

One of the things we can learn by messing around with Lorentz transformations is that $\sigma^2\psi_R^*$ transforms as a left-handed spinor despite being constructed out of ψ_R. This is a conclusion that only works when there is no explicit mass term in the Lagrangian. Otherwise it would make no sense to try to separate ψ_R from ψ_L! Let us therefore write, suppressing generation indices for now,

$$\begin{aligned} \tilde{\psi}_L &= \sigma^2\psi_R^* \\ \tilde{\psi}_L^\dagger &= \psi_R^T\sigma^2 \end{aligned} \tag{725}$$

so that

$$\begin{aligned} \psi_R &= -\sigma^2\tilde{\psi}_L^* \\ \psi_R^\dagger &= -\tilde{\psi}_L^T\sigma^2 \end{aligned} \tag{726}$$

Let us examine the free and minimal-coupling bits separately. We will write

$$S_R = \int \psi_R^\dagger i\sigma^\mu \partial_\mu \psi_R = i\int \psi_{R\,\alpha}^\dagger \sigma^\mu_{\alpha\beta}\vec{\partial}_\mu \psi_{R\,\beta} = -i\int \psi_{R\,\alpha}^\dagger \sigma^\mu_{\alpha\beta}\overleftarrow{\partial}_\mu \psi_{R\,\beta} \tag{727}$$

where in the last step we used integration by parts. Next, we use the fact that animals like $\psi_{R\,\beta}$ or $\sigma^\mu_{\alpha\beta}$ is a number, not a vector or a matrix in Dirac index space or any other space. The only pieces we have to be careful about commuting through each other merrily are the Grassmann variables. Therefore,

$$\begin{aligned} S_R &= +i\int \psi_{R\,\beta}\sigma^\mu_{\alpha\beta}\vec{\partial}_\mu \psi_{R\,\alpha}^\dagger \\ &= +i\int (-\sigma^2_{\beta\gamma}\tilde{\psi}_L^\dagger)\sigma^\mu_{\alpha\beta}\partial_\mu(-\tilde{\psi}^T_{L\,\delta}\sigma^2_{\delta\alpha}) \\ &= +i\int \tilde{\psi}^\dagger_{L\,\gamma}\sigma^2_{\beta\gamma}\sigma^\mu_{\alpha\beta}\sigma^2_{\delta\alpha}\partial_\mu\tilde{\psi}_{L\,\delta} \\ &= +i\int \tilde{\psi}^\dagger_{L\,\gamma}(\sigma^{2,T})_{\gamma\beta}(\sigma^{\mu,T})_{\beta\alpha}(\sigma^{2,T})_{\alpha\delta}\partial_\mu\tilde{\psi}_{L\,\delta} \end{aligned} \tag{728}$$

But since $\sigma^{i\,\dagger} = \sigma^i$ we get $(\sigma^i)^T = (\sigma^i)^* = \sigma^2\sigma^i\sigma^2$, we can rewrite the action for right-handed fermions as

$$\begin{aligned} S_R &= +i\int \tilde{\psi}_L^\dagger(\sigma^2\sigma^2\sigma^2)(\sigma^2\bar{\sigma}^\mu\sigma^2)(\sigma^2\sigma^2\sigma^2)\partial_\mu\tilde{\psi}_L \\ &= \int \tilde{\psi}_l^\dagger i\bar{\sigma}^\mu\partial_\mu\tilde{\psi}_L \end{aligned} \tag{729}$$

Note that this is exactly what we hoped for: the expected action principle for a left-handed fermion $\tilde{\psi}_L$.

Now let us look at the vertex term and make all indices explicit to keep the algebra straight:

$$\begin{aligned} S_{\rm int} &= \int g\psi_R^\dagger\sigma^\mu A^A_\mu t^A_r\psi_R \\ &= \int g(\psi_R^\dagger)_{\alpha a}\sigma^\mu_{\alpha\beta}A^A_\mu(t^A_r)_{ab}(\psi_R)_{\beta b} \\ &= \int g(-\tilde{\psi}_L^T\sigma^2)_{\alpha a}\sigma^\mu_{\alpha\beta}A^A_\mu(t^A_r)_{ab}(-\sigma^2\tilde{\psi}_L^*)_{\beta b} \\ &= \int g\tilde{\psi}^\dagger_{L\gamma_a}(\sigma^2)_{\gamma\alpha}(\sigma^\mu)_{\alpha\beta}A^A_\mu(t^A_r)_{ab}(\sigma^2)_{\beta\delta}\tilde{\psi}^*_{L\delta b} \end{aligned} \tag{730}$$

Now that all of our indices are explicit, we can start the commute-it-through operation while remembering carefully the rules for anticommuting Grassmann fields. We get

$$\begin{aligned} S_{\rm int} &= -\int g\tilde{\psi}^*_{L\delta b}(\sigma^2)_{\beta\delta}\sigma^\mu_{\alpha\beta}A^A_\mu(t^A_r)_{ab}(\sigma^2)_{\gamma\alpha}\tilde{\psi}^*_{L\gamma a} \\ &= -\int g\tilde{\psi}^\dagger_{L\delta b}(\sigma^{2,T})_{\delta\beta}(\sigma^{\mu,T})_{\beta\alpha}A^A_\mu(t^{A,T}_r)_{ba}(\sigma^{2,T})_{\alpha\gamma}\tilde{\psi}_{L\gamma a} \end{aligned} \tag{731}$$

Now, using the fact that $\sigma^2\sigma^\mu\sigma^2 = \bar{\sigma}^\mu$ and vice versa, and the complex conjugation property of sigma matrices,

$$S_{\rm int} \;=\; -\int g\tilde{\psi}^\dagger_{\delta b}(\sigma^2)_{\delta\beta}(\sigma^{\mu *})_{\beta\alpha}(\sigma^2)_{\alpha\gamma}A^A_\mu(t^{A,T}_r)_{ba}\tilde{\psi}_{\gamma a} \tag{732}$$

In other words, we have

$$S_{\rm int} = -\int g\tilde{\psi}^\dagger_L\bar{\sigma}^\mu A^A_\mu(t^A_r)^T\tilde{\psi}_L \tag{733}$$

This shows (without using any integration by parts manoeuvre) that the representation $\tilde{\psi}_L$ has generators $(t^A_r)^T$ if ψ_R had generators (t^A_r). It follows immediately that our twiddled fields belong to the *conjugate* representation of the gauge group. This is known as the $\bar{\mathbf{r}}$ if the original representation is labelled by $\mathbf{r}$.

Putting the right- and left-handed actions together gives, for one species,

$$\begin{aligned}
\mathscr{L} &= \psi^\dagger_R i\sigma^\mu\left(\mathbb{1}\partial_\mu - igA^A_\mu t^A_r\right)\psi_R \\
&= \tilde{\psi}^\dagger_L i\bar{\sigma}^\mu\left(\mathbb{1}\partial_\mu + igA^A_\mu (t^A_r)^T\right)\tilde{\psi}_L \\
&= \tilde{\psi}^\dagger_L i\bar{\sigma}^\mu\left(\mathbb{1}\partial_\mu - igA^A_\mu (t^A_{\bar{r}})\right)\tilde{\psi}_L
\end{aligned} \tag{734}$$

Consider QCD with n_f flavours of massless quarks. This can be rewritten as an $SU(3)$ gauge theory coupled to n_f massless quarks in the $\mathbf{3}$ representation and n_f massless quarks in the $\bar{\mathbf{3}}$ representation.

How about the most general gauge theory with massless fermions? Let us put the left-handed fermions into an arbitrary, *reducible* representation R of the gauge group G. We just found right now that rewriting a system of (massless) Dirac fermions in a solely-left-handed basis gives $\mathbf{R} = \mathbf{r}\oplus\bar{\mathbf{r}}$. Note: $\bar{\mathbf{r}}$ may be equivalent to $\mathbf{r}$ if there exists a unitary U such that $t^A_{\bar{r}} = -t^{A*}_r = -(t^A_r)^T$. If such a U exists, then the representation $\mathbf{r}$ is *real*. Conversely, if R is not a real representation, then $\mathscr{L}$ cannot be rewritten in terms of Dirac fermions and is intrinsically chiral.

At the classical level, R is unrestricted. At loop level, for chiral gauge theories there will typically be anomalies which render many of the field theories inconsistent quantum mechanically. Let us compute which Feynman diagrams would contribute. We have identical-looking Feynman diagrams to the two we summed in the previous section. In particular, we had

- The marked vertex denoted by $\bullet$ has one insertion of the gauge symmetry current

$$g^{\mu A} = \bar{\psi}\gamma^\mu\left(\frac{\mathbb{1}-\gamma_5}{2}\right)t^A\psi \tag{735}$$

- Gauge boson vertices also carry a left-chiral projector $\mathscr{P}_- = (\mathbb{1}-\gamma_5)/2$. The three projectors emanating from the gauge vertices or symmetry current insertions in the 1-loop diagram can be collected into one, since $\mathscr{P}^2_\pm = \mathscr{P}_\pm$.
- The technique we used earlier – regularizing with $l_\parallel$ and $l_\perp$ – yields for the term in γ_5 an *axial vector anomaly*

$$\langle p,\nu,b;k,\lambda,c|\partial_\mu j^{\mu 5}|0\rangle = \frac{g^2}{8\pi^2}\epsilon^{\alpha\nu\beta\lambda}p_\alpha k_\beta\mathscr{A}^{ABC} \tag{736}$$

where the group theory factor is

$$\mathscr{A}^{ABC} = \mathrm{Tr}\left[t^A\left\{t^B, t^C\right\}\right] \tag{737}$$

Therefore,

$$\partial_\mu j^{\mu 5} = 0 \quad \text{iff} \quad \mathscr{A}^{ABC} = 0 \tag{738}$$

Note that this $\mathscr{A}^{ABC}$ is a group invariant which is totally symmetric in all three indices. Therefore, it doesn't actually matter which gauge boson leg we think of as being 'attached' to the current insertion in the triangle anomaly diagram.

Local gauge symmetry of the action relies on the same *global* symmetry being exact. So if the chiral current is not conserved, then our gauge field theory is, well, screwed. Triangle anomaly diagrams like the above typically generate divergent mass terms for non-Abelian gauge bosons, which messes with the delicate relationships between 3- and 4-point vertices (Ward Identities, etc.). So unitarity of the gauge theory S-matrix demands the consistency condition

$$\mathscr{A}^{ABC} = 0 \tag{739}$$

Gauge field theories involving chiral couplings to matter are said to be anomaly-free if they obey this equation.

Note: if ψ are real then the anomaly coefficients vanish. A special case of this is the Dirac representation $\mathbf{R} = \mathbf{r} \oplus \bar{\mathbf{r}}$; all theories built with Dirac fields are automatically anomaly-free.

Let's do some examples.

$SU(2)^3$: For this case, $t^A = \sigma^A/2 \equiv \tau^A$ and we have $\{\sigma^B, \sigma^C\} = 2\delta^{BC}$ so that

$$\mathscr{A}^{ABC} = \frac{1}{8}\mathrm{Tr}[\sigma^A \cdot 2\delta^{BC}] = 0 \tag{740}$$

$SU(2)^2 U(1)^1$: For this case we have

$$\mathscr{A}^{BC} = \mathrm{Tr}[Q\{\tau^B, \tau^C\}] = \frac{1}{2}\mathrm{Tr}[Q]\delta^{BC} \tag{741}$$

How can the trace of the charge operator vanish? It can happen if we remember that we have to sum over all quarks and leptons! For quarks and leptons of the Standard Model we have

$$\mathrm{Tr}[Q] = 3 \cdot (+\frac{2}{3} - \frac{1}{3}) + (0 - 1) = 1 - 1 = 0 \tag{742}$$

where the factor of 3 in the first term is just the number of colours. So we see that the electroweak theory of the standard model is consistent *only* when the number of leptons is equal to the number of quarks. You can also check for yourself by working out the explicit details that the Standard Model is consistent only if quarks and leptons come in complete $SU(2)$ doublets: it is not theoretically possible to have an *odd* number of quarks or leptons. This is one reason why everyone expected the top quark after the discovery of the bottom quark: it just "had to be there"!

One interesting question to ask is: which gauge groups possess anomaly coefficients in the first place? These things $\mathscr{A}^{ABC}$ are completely symmetric tensors which are also group invariants. In some cases the structure of the group precludes there being anomaly coefficients at all. For $SU(2)$, we can ask what happens when we tensor together two spin-1 representations. We get the direct sum of a singlet and a spin-2 rep. Neither of these would couple with a spin-1 rep to give a group invariant. The conclusion is that $SU(2)$ has no anomaly coefficient.

For $SU(n), n \geq 3$, there is such a group invariant. It arises from the anticommutation relations of generators

$$\{t_n^A, t_n^B\} = \frac{1}{n}\delta^{AB} + d^{ABC} t_n^C \tag{743}$$

The unique group invariant is the d^{ABC} thingy.

Let us define a scalar anomaly coefficient $\mathscr{A}$ by

$$\mathrm{Tr}\left[t_r^A \left\{t_r^B, t_r^C\right\}\right] = \frac{1}{2}\mathscr{A}(r) d^{ABC} \tag{744}$$

For the fundamental representation $\mathbf{n}$, we have $A(\mathbf{n}) = 1$.

Of the various available simple Lie groups, only $SU(n)$, $SO(4n+2)$ and E_6 have complex representations – the kind we need to represent chiral fermions. Of these, only $SU(n)$ and $SO(6) \simeq SU(4)$ have the required three-index tensor group invariant to form an anomaly. Gauge theories based on $SO(4n+2), n \geq 2$ or on E_6 are anomaly-free. This fact is related to the popularity of $SO(10)$ and E_6 in GUT model building!

12 Appendix: advanced tidbits

12.1 YM vs GR

The story of Yang-Mills may appear to possess some strong similarities to the story of the Christoffel connection or spin connection in GR. There are however very important physical differences, chief among them that GR is the classical theory for a massless spin-two field, while YM is for spin one. In addition, general coordinate transformations act on spacetime coordinates whereas gauge transformations act on an *internal* field space.

There are also crucial mathematical differences between YM and GR. The gauge connection is technically defined as a connection on a principal fibre bundle, whereas the spin connection (not to be confused with the Christoffel connection) is a connection on a bundle of orthonormal frames.

In GR, we can define our dynamical gravitational fields to be vielbeins e^a_μ via

$$g_{\mu\nu} = e^a_\mu e^b_\nu \eta_{ab} \tag{745}$$

Physicists will sometimes play fast and loose with words by calling vielbeins the square root of the metric tensor. What is most physically crucial about these vielbein animals is that they convert curved spacetime indices (μ) to flat tangent space indices (a). The fact that curved and tangent space indices are not equivalent is because we allow a nontrivial spacetime metric; in simple Minkowski space they would be the same. Curved indices are raised or lowered by the full metric tensor, while tangent space indices are raised or lowered via the Minkowski metric.

It is possible to define tangent-space one-forms (one-index covariant vectors) via

$$e^a = e^a_\mu dx^\mu \tag{746}$$

Notice that what we did here was to contract up the spacetime (curved) index of the vielbein with a coordinate basis one-form dx^μ to create a one-form living in the tangent space. This guy would then be subject to Lorentz transformations, while the curved-space counterpart suffers full coordinate transformations.

The requirement that the spin connection be torsion-free (which is what GR assumes) is

$$T^a = de^a + \omega^a{}_b \wedge e^b = 0 \tag{747}$$

This equation is sufficient to determine $\omega^a{}_b$ in terms of vielbeins and their first derivatives. The spin connection coefficients obey

$$\omega_{ab} = -\omega_{ba} \tag{748}$$

Then the *Cartan structure equations* read

$$R^a{}_b = d\omega^a{}_b + \omega^a{}_c \wedge \omega^c{}_b \tag{749}$$

Note that the Riemann tensor's indices in curved space would be recovered by using

$$R^a{}_b = R^a{}_{b\mu\nu} dx^\mu \wedge dx^\nu \tag{750}$$

and contracting with (inverse) vielbeins to convert from tangent space to curved space

$$R^{\mu}{}_{\nu\lambda\sigma} = e^{\mu}_{a} e^{b}_{\nu} R^{a}{}_{b\lambda\sigma} \tag{751}$$

As you can see, the structure equation (749) for the spin connection bears a strong similarity to how the field strength F is defined in terms of the gauge potential A:

$$F = dA - igA \wedge A \tag{752}$$

This similarity does not persist at a deeper level in generic dimensions of spacetime. Gauge and gravitational fields really are physically (and mathematically!) different. Insisting otherwise amounts to wishful thinking technically.

12.2 Conformal group

The conformal group $SO(d+1,2)$ is the subgroup of general coordinate transformations preserving the conformal flatness of the Minkowski metric. It contains the Lorentz group $SO(d,1)$ as a subgroup. It also contains two new kinds of transformations. The first kind is scale transformations a.k.a. dilatations, which act as

$$x^{\mu} \to x'^{\mu} = (1+\lambda) x^{\mu} \,, \tag{753}$$

where λ is a constant parameter. The second new kind of transformation is special conformal transformations a.k.a. conformal boosts. There is a simple way of deriving the effect of a conformal boost: do an inversion

$$x^{\mu} \to x'^{\mu} = \frac{x^{\mu}}{x^2} \tag{754}$$

followed by a translation followed by another inversion. To show that dilatations and inversions preserve the conformal flatness of the metric, we can inspect how $ds^2 = \eta_{\mu\nu} dx^{\mu} dx^{\nu}$ transforms.

Using the inversion-translation-inversion idea, we can show that the conformal boosts act as

$$x^{\mu} \to \frac{x^{\mu} + b^{\mu} x^2}{1 + 2b \cdot x + b^2 x^2} \,, \tag{755}$$

where b^{μ} is a constant parameter. It is straightforward to use this to find the infinitesimal form of the conformal boosts. For a Lorentz transformation of the form

$$U = \exp\left(\frac{i}{2} \omega^{\mu\nu} M_{\mu\nu} \right) \tag{756}$$

acting on the coordinates, the infinitesimal form is

$$\Delta x^{\mu} = \omega^{\mu\nu} x_{\nu} \,. \tag{757}$$

Using the above facts, we can find the most general infinitesimal conformal transformation,

$$\Delta x^{\mu} = a^{\mu} + \omega^{\mu}{}_{\nu} x^{\nu} + \lambda x^{\mu} + b^{\mu} x^2 - 2x^{\mu} (b \cdot x) \,, \tag{758}$$

where $\lambda, a^\mu, b^\mu, \omega^\mu{}_\nu$ are constant parameters corresponding to dilatations, translations, conformal boosts, and Lorentz transformations respectively. Counting the number of parameters for the conformal group, we can quickly see that there are $n(D) = \frac{1}{2}(D+2)(D+1)$ of them.

It is interesting to work out the algebra of the generators of infinitesimal conformal symmetries, which contains the Poincaré algebra of HW1 as a subalgebra. Using the representations[11]

$$\begin{aligned}
P_\mu &= -i\partial_\mu\,, \\
L_{\mu\nu} &= -x_\mu P_\nu + x_\nu P_\mu\,, \\
D &= -ix^\mu \partial_\mu\,, \\
K_\mu &= i(x^2\partial_\mu - 2x_\mu x^\nu \partial_\nu)\,.
\end{aligned} \tag{759}$$

Using our knowledge from HW1, we can find the conformal algebra commutation relations,

$$\begin{aligned}
[D, P_\mu] &= +iP_\mu\,, \\
[D, K_\mu] &= -iK_\mu\,, \\
[K_\mu, P_\nu] &= +2i(\eta_{\mu\nu} D - L_{\mu\nu})\,, \\
[P_\rho, L_{\mu\nu}] &= +i(\eta_{\rho\mu} P_\nu - \eta_{\rho\nu} P_\mu)\,, \\
[K_\rho, L_{\mu\nu}] &= +i(\eta_{\rho\mu} K_\nu - \eta_{\rho\nu} K_\mu)\,, \\
[L_{\mu\nu}, L_{\rho\sigma}] &= +i(\eta_{\nu\rho} L_{\mu\sigma} - \eta_{\mu\rho} L_{\nu\sigma} + \eta_{\mu\sigma} L_{\nu\rho} - \eta_{\nu\sigma} L_{\mu\rho})\,.
\end{aligned} \tag{760}$$

Making the definitions

$$J_{\mu\nu} = L_{\mu\nu}\,, \qquad J_{0\mu} = \frac{1}{2}\left(P_\mu + K_\mu\right)\,, \qquad J_{-1\mu} = \frac{1}{2}\left(P_\mu - K_\mu\right)\,, \qquad J_{-10} = D\,, \tag{761}$$

allows us to obtain the $SO(d+1,2)$ commutation relations,

$$[J_{AB}, J_{CD}] = +i\left(\eta_{AD} J_{BC} - \eta_{AC} J_{BD} + \eta_{BC} J_{AD} - \eta_{BD} J_{AC}\right)\,, \tag{762}$$

where $J_{AB} = -J_{BA}$, (η_{AB}) =diag$(-1,-1,1,\ldots,1)$, and A, B range over $-1, 0, 1, \ldots, d$.

The two-dimensional case is special because the conformal algebra we have been discussing above turns out to be just a subalgebra of an *infinite*-dimensional symmetry algebra. In 2D, conformal transformations are analytic coordinate transformations

$$z \to f(z)\,, \qquad \bar{z} \to \bar{f}(\bar{z})\,. \tag{763}$$

To exhibit the generators, consider infinitesimal conformal transformations of the form

$$z \to z' = z - \epsilon_n z^{n+1} \qquad \bar{z} \to \bar{z}' = \bar{z} - \bar{\epsilon}_n \bar{z}^{n+1} \tag{764}$$

where $n \in \mathbb{Z}$. These are angle-preserving whenever f and its inverse are holomorphic. The corresponding infinitesimal generators are

$$\ell_n = -z^{n+1}\partial \qquad \bar{\ell}_n = -\bar{z}^{n+1}\bar{\partial} \tag{765}$$

[11] We use conventions of the conformal field theory textbook by di Francesco Mathieu and Senechal.

where we use the shorthand $\partial = \partial/\partial z$ and $\bar{\partial} = \partial/\partial\bar{z}$. By considering the infinitesimal transformations generated by the $\{\ell_{0,\pm1}, \bar{\ell}_{0,\pm1}\}$ only, it is straightforward to show that ℓ_{-1} and $\bar{\ell}_{-1}$ generate translations, $(\ell_0 + \bar{\ell}_0)$ generates scalings, $i(\ell_0 - \bar{\ell}_0)$ generates rotations, and $\ell_1, \bar{\ell}_1$ generate special conformal transformations. Also, the $\ell_n, \bar{\ell}_n$ satisfy the classical Virasoro algebras:

$$[\ell_m, \ell_n] = (m-n)\ell_{m+n}\,, \qquad [\bar{\ell}_m, \bar{\ell}_n] = (m-n)\bar{\ell}_{m+n}\,. \tag{766}$$